A. Supplementary Material

This is the supplementary material for paper *What Is the Optimal Ranking Score Between Precision and Recall? We Can Always Find It and It Is Rarely* F_1.

Contents

A.1. Supplementary Material About Sec. 2.2.(Building Upon the Related Work)

A.1.1. Wide Use of F_1

The F_1 score (*a.k.a.* Dice) is definitely standard in computer vision, serving as ranking criteria in benchmarks and challenges on tasks such as image/video classification in medical imaging (COVID-CT, PAIP, ECDP, TN-SCUI), and behavior analysis (ABAW), segmentation in video surveillance (IWDD [2]), medical imaging (3D-TEETH-SEG, DRAC, UNICORN), and visual quality control (VAND), detection in medical imaging (MIDOG, UNICORN). Also, CADA-RRE [17] uses F_2; we discuss it in Appendix A.2.4. Notably, some are from CVPR'25 (VAND, ABAW). More exist in computer vision and beyond.

A.1.2. Reminder of the Three Axioms of Performance-Based Ranking

These axioms are from last year CVPR paper [31].

The first axiom states that a ranking must be established based on a preorder $\lesssim$ on the performances. Following Theorem 1 of [31], in this paper, we consider only preorders $\lesssim_X$ induced by a score X in the following manner. Regardless of whether the performances P_1 and P_2 to be compared belong to the domain of definition of X, we decide that $P_1 \lesssim_X P_2$ when $P_1 = P_2$. If P_1 and P_2 are both in the domain of definition of X, then we decide that $P_1 \lesssim_X P_2$ if and only if $X(P_1) \leq X(P_2)$. In all other cases, $P_1 \not\lesssim_X P_2$. Such a preorder is called *performance ordering induced by* X. As an immediate consequence of this first axiom, the rankings are stable: it is impossible to swap two previously compared methods by inserting or deleting methods in the ranking.

The two other axioms ensure that one can interpret the binary relation $\lesssim_X$ as *worse or equivalent*. For the sake of consistency, all the other binary relations of interest (*worse than*, *better than*, *equivalent to*, *incomparable with*, *etc.*) are derived from $\lesssim_X$. When the two following axioms are satisfied, we say that the performance ordering induced by X satisfy the axioms of performance-based ranking, or more simply that X satisfy the axioms.

The second axiom specifies a condition that must be satisfied for the preorder $\lesssim_X$ on performances to be compatible with the task. In the case of the two-class crisp classification task studied in this paper, the satisfaction S is binary: we are not at all satisfied with a false positive or a false negative, and we are fully satisfied with a true negative and a true positive. Denoting the accuracy by A, the second axiom, given in a generic form in [31], is in this case equivalent to three conditions: (1) there is no performance worse than a performance P such that $A(P) = 0$; (2) if a performance P_1 is such that $A(P_1) = 0$ and a performance P_2 is such that $A(P_2) = 1$, then P_1 can never be better than P_2; (3) there exists no performance better than a performance P such that $A(P) = 1$. The second axiom thus focuses on the extreme cases of the worst and best performances, but says nothing about intermediate performances. It should be noted that it is very permissive. A score may put performances for which $A(P) > 0$ on the same footing the worst performances, or performances for which $A(P) < 0$ on the same footing the best performances. This is for example the case of F_1: $A = 0 \Rightarrow F_1 = 0$ but $F_1 = 0 \not\Rightarrow A = 0$.

The third axiom specifies a condition that must be satisfied for the preorder $\lesssim_X$ on performances to be compatible with evaluation (*i.e.*, the mapping from methods to performances). It is this third axiom that constrains the ordering of performances that are neither among the best nor among the worst ones. Consider a *blind* and non-deterministic combination of some base methods to form a hybrid method: each time the hybrid method is used, it begins by randomly selecting one of the base methods, *independently of its input*, and then executes that method. As the combination is *blind*, it would not make sense for the hybrid method to be better or worse than the best or the worst of the combined methods, respectively. Note that this axiom is not satisfied by $SIVF$ in general, but is satisfied by it when the priors are fixed.

A.1.3. All F_β Lead to Meaningful Performance Orderings

It has been shown in [31] that all performance orderings derived from some score of the form

$$R_I(P) = \frac{I(tn)P(\{tn\}) + I(tp)P(\{tp\})}{I(tn)P(\{tn\}) + I(fp)P(\{fp\}) + I(fn)P(\{fn\}) + I(tp)P(\{tp\})}, \tag{23}$$

where I denotes a positive random variable called *Importance*, satisfy all the axioms of the theory of performance-based ranking. These scores are known as *ranking scores*.

In order to prove that all F_β lead to meaningful performance orderings, we show here that the F_β are particular cases of R_I. The proofs for $\beta = 1/2$, $\beta = 1$, and $\beta = 2$ were already given in [31], so here we give a generalization. Our proof is based on the well known fact that F_β has two nearly equivalent definitions. The first one, that we have used in the introduction (Eq. (1)), is as follows:

$$F_\beta : \left\{P \in \mathbb{P}_{(\Omega,\Sigma)} : P(\{tp\}) \neq 0\right\} \to [0,1] : P \mapsto \left(\frac{1}{1+\beta^2} Pr^{-1}(P) + \frac{\beta^2}{1+\beta^2} Re^{-1}(P)\right)^{-1}. \tag{24}$$

The second one is

$$F_\beta : \left\{P \in \mathbb{P}_{(\Omega,\Sigma)} : P(\{tn\}) \neq 1\right\} \to [0,1] : P \mapsto \frac{(1+\beta^2)\,P(\{tp\})}{1\,P(\{fp\}) + \beta^2\,P(\{fn\}) + (1+\beta^2)\,P(\{tp\})}\,. \tag{25}$$

It turns out that the first definition is a restriction of the second one: $P(\{tp\}) \neq 0 \Rightarrow P(\{tn\}) \neq 1$ and both F_β are equal on $\left\{P \in \mathbb{P}_{(\Omega,\Sigma)} : P(\{tp\}) \neq 0\right\}$. It is now possible to compare Eq. (25) with Eq. (23) and to see that, for any given β, F_β is a particular case of ranking score with

$$(I(tn), I(fp), I(fn), I(tp)) \propto \left(0, 1, \beta^2, 1+\beta^2\right)\,, \tag{26}$$

or equivalently

$$(I(tn), I(fp), I(fn), I(tp)) \propto (0, 1-b, b, 1)\,, \tag{27}$$

with $b = \beta^2/1+\beta^2$.

A.1.4. $SIVF$ Leads to a Meaningful Performance Ordering When the Class Priors Are Fixed

The skew-insensitive version of F_1 as been defined in [10] as

$$SIVF = \frac{2TPR}{TPR + FPR + 1}\,. \tag{28}$$

When the class priors are fixed and such that $\pi_- \neq 0$ and $\pi_+ \neq 0$,

$$\begin{aligned} SIVF(P) &= \frac{2\frac{P(\{tp\})}{\pi_+}}{\frac{P(\{tp\})}{\pi_+} + \frac{P(\{fp\})}{\pi_-} + 1} = \frac{2\,\pi_-\,P(\{tp\})}{\pi_-\,P(\{tp\}) + \pi_+\,P(\{fp\}) + \pi_-\,\pi_+} \\ &= \frac{2\,\pi_-\,P(\{tp\})}{\pi_-\,P(\{tp\}) + \pi_+\,P(\{fp\}) + \pi_-\,(P(\{fn\}) + P(\{tp\}))} = \frac{2\,\pi_-\,P(\{tp\})}{\pi_+\,P(\{fp\}) + \pi_-\,P(\{fn\}) + 2\,\pi_-\,P(\{tp\})}\,. \end{aligned} \tag{29}$$

The comparison of Eq. (29) with Eq. (23) shows that $SIVF$ is a particular case of ranking score with

$$(I(tn), I(fp), I(fn), I(tp)) \propto \left(0, \pi_+, \pi_-, 2\,\pi_-\right)\,. \tag{30}$$

A.1.5. $SIVF$ Leads to the Same Ranking as F_β with $\beta^2 = \frac{\pi_-}{\pi_+}$

A first demonstration, based on the relationship between $SIVF$ and F_β. From [10], it is already known that $SIVF = F_1$ when $\pi_- = \pi_+ = 1/2$. Here, we give a generalization for $\pi_- \neq 0$ and $\pi_+ \neq 0$. We noticed that

$$\beta^2 = \frac{\pi_-}{\pi_+} \Rightarrow F_\beta = \frac{SIVF}{(\pi_+ - \pi_-)\,SIVF + 2\pi_-}\,. \tag{31}$$

When $\pi_- \neq 0$, we have

$$\beta^2 = \frac{\pi_-}{\pi_+} \Rightarrow \frac{\partial F_\beta}{\partial SIVF} > 0\,. \tag{32}$$

Therefore, when the priors are fixed, $SIVF$ leads to the same performance ordering as F_β with $\beta^2 = \frac{\pi_-}{\pi_+}$.

A second demonstration, based on the importance values. Property 4 of [31], when particularized for two-class classification, states that the performance ordering induced by a ranking score is insensitive to the uniform scaling of the importance given to the unsatisfying samples (fp and fn) or to the satisfying (tn and tp) samples. The comparison between Eq. (26) and Eq. (30) shows that:

- for the unsatisfying samples, we have $\left(\bullet, 1, \beta^2, \bullet\right) \propto \left(\bullet, \pi_+, \pi_-, \bullet\right)$ when $\beta^2 = \frac{\pi_-}{\pi_+}$;
- for the satisfying samples, we have always $\left(0, \bullet, \bullet, 1+\beta^2\right) \propto \left(0, \bullet, \bullet, 2\,\pi_-\right)$.

Therefore, when the priors are fixed, $SIVF$ leads to the same performance ordering as F_β with $\beta^2 = \frac{\pi_-}{\pi_+}$.

A.2. Supplementary Material About Sec. 3 (Theory)

A.2.1. On the Minimization of Fréchet Variance (Eq. (8))

In Sec. 3.4.1, we say that the optimal tradeoffs are the F_β scores that minimize the Fréchet variance [11]:

$$\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)\,.$$

We also say that the solutions, known as the Karcher means [21] are those that are equidistant of Pr and Re, *i.e.* such that

$$\begin{aligned} & d_\tau(Pr; F_*) = d_\tau(F_*; Re) = \frac{d_\tau(Pr; Re)}{2} \\ \Leftrightarrow\ & \tau(Pr; F_*) = \tau(F_*; Re) = \frac{1 + \tau(Pr; Re)}{2}\,. \end{aligned}$$

Hereafter, we discuss here the case of a continuum of rankings, as shown in Fig. A.2.1a. The case of a finite amount of rankings is depicted in Fig. A.2.1b and Fig. A.2.1c for, respectively, an odd and an even amount.

By restricting to the family of F_β scores, the solution is actually unique and is thus a Fréchet mean. Indeed, since Eq. (6) states that

$$d_\tau(Pr; Re) = d_\tau(Pr; F_\beta) + d_\tau(F_\beta; Re) \qquad \forall \beta \geq 0$$

we have

$$\begin{aligned} \sigma^2(\beta) &= d_\tau^2(Pr; F_\beta) + (d_\tau(Pr; Re) - d_\tau(Pr; F_\beta))^2 && (33) \\ &= 2d_\tau^2(Pr; F_\beta) - 2d_\tau(Pr; Re)d_\tau(Pr; F_\beta) + d_\tau^2(Pr; Re) && (34) \end{aligned}$$

which is quadratic in $d_\tau(Pr; F_\beta)$ and thus uniquely minimized for β such that

$$d_\tau(Pr; F_\beta) = \frac{d_\tau(Pr; Re)}{2}\,. \tag{35}$$

This incidentally yields the same equality for $d_\tau(F_\beta; Re)$. The result in terms of correlation stems from the relation $\tau = 1 - 2d_\tau$.

A.2.2. On the Close-Form Expression for the Optimal β (Eqs. (11) and (12))

Complexity/scalability. After evaluating n classifiers, a straightforward computation of the optimal β with our formulas requires indeed $O(n^2)$ uses of Eq. (11) (same as for Kendall's τ). Even for $n = 1000$ (hardly met in practice), this takes less than one second on a modern laptop.

An alternative to Eq. (11). We thank Peter Flach for pointing out that Eq. (11) can be rewritten in terms of only precision and recall values as follows:

$$\vartheta(P_1, P_2) = -\frac{Pr^{-1}(P_1) - Pr^{-1}(P_2)}{Re^{-1}(P_1) - Re^{-1}(P_2)}\,. \tag{36}$$

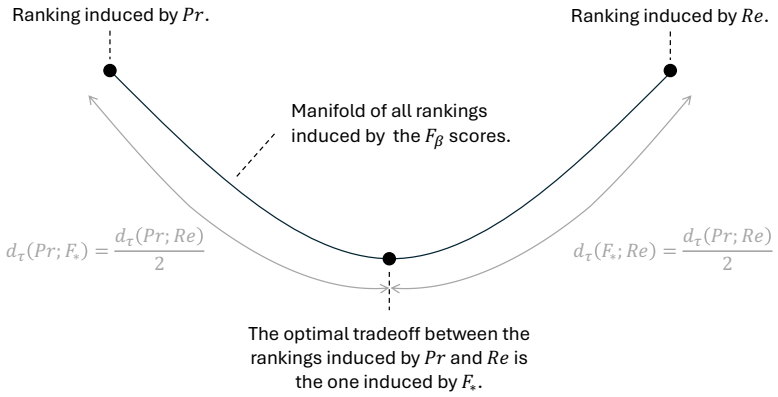
This formula is Eq. (6) in [9].

Downstream objective. We can generalize Eq. (8) by weighting d_τ^2 and Eq. (12) by replacing the median by a quantile. To ensure that the 0% quantile corresponds to Pr and that the 100% quantile corresponds to Re, the quantile is taken over $\{0, \infty\} \cup \{\vartheta(P_i, P_j) \,|\, i \neq j \wedge \vartheta(P_i, P_j) \geq 0\}$. For instance, to weight four times more Re than Pr, from the *point of view of values* one uses $b = 0.8$ in Eq. (1) (*i.e.*, $\beta = 2$, as in CADA-RRE). But from the *point of view of ranks*, we take the 80^{th} percentile of ϑ values. This leads to $\beta = 0.914$ in the example of CADA-RRE (Appendix A.2.4). The relationship between the downstream objective (*i.e.*, the chosen quantile) and the optimal value for β is shown in Fig. A.2.2.

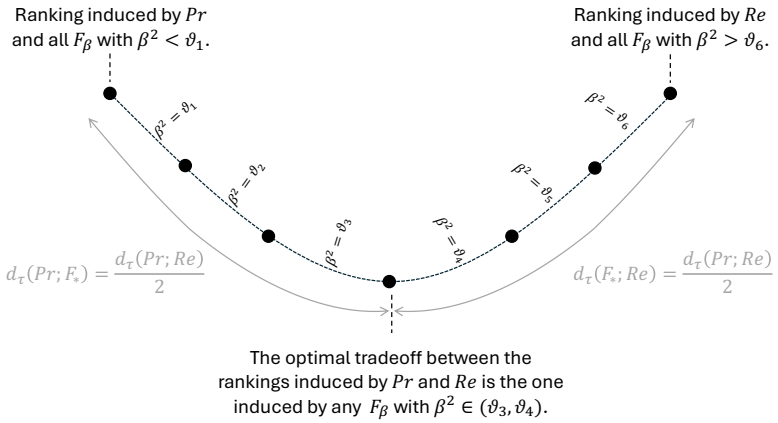
A.2.3. On the Formulas for Degree of Optimality of Some F_β

We provide here a few explanations about how Eqs. (13) to (15) cen be derived. For that, let us remind three interpretations for Kendall's distance $d_\tau(X_1; X_2) \in [0, 1]$.

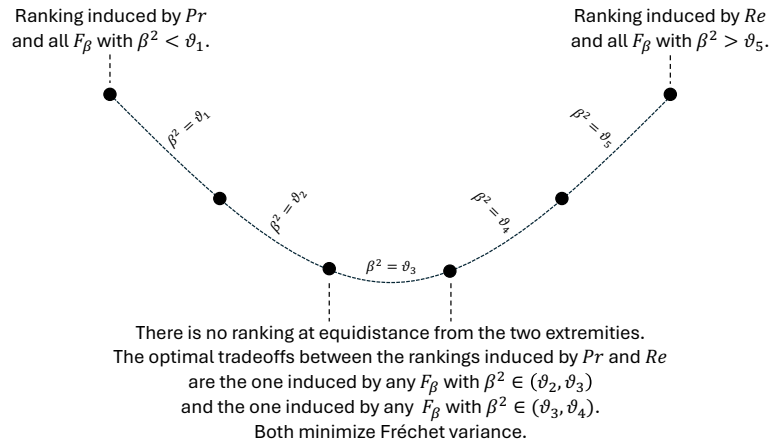
- First, Eq. (5) shows that $d_\tau(X_1; X_2)$ is equal to the proportion of pairs of elements for which X_1 and X_2 do not agree on the relative order. This interpretation is the cornerstone for probabilistic reasoning.



(a) **Case in which there is a bijection between the values of β and the rankings induced by F_β.** The rankings form a continuum (manifold) and the minimization of Fréchet variance leads to the midpoint where $d_\tau(Pr; F_*) = d_\tau(F_*; Re)$.



(b) **Case in which the number of rankings induced by the F_β scores is odd.** The rankings form a path graph. The nodes correspond to the rankings, each of them being induced by a range of values for β. The edges correspond to the swaps that occur between the consecutive rankings, at some given values ϑ for the β parameter. Assuming there is no group of at least three co-aligned performances, the median of all these ϑ values belongs to the range of βs for which the ranking is at the midpoint where $d_\tau(Pr; F_*) = d_\tau(F_*; Re)$.



(c) **Case in which the number of rankings induced by the F_β scores is even.** Assuming there is no group of at least three co-aligned performances, the median of all the ϑ values corresponds to the swap between the two rankings minimizing Fréchet variance.

Figure A.2.1. The optimal tradeoff as the solution of the minimization of Fréchet variance $\sigma^2(\beta)$.

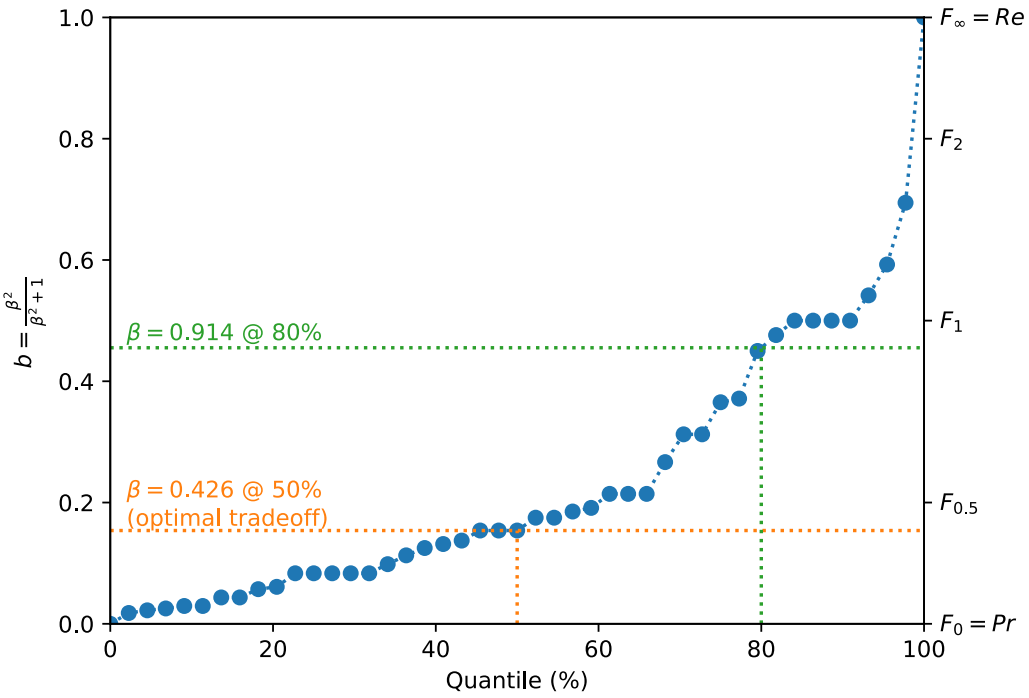


Figure A.2.2. The relationship between the downstream objective (*i.e.*, the chosen quantile) and the optimal value for β, for the CADA-RRE example described in Appendix A.2.4. The points correspond to the various values of β for which a swap between two consecutive classifiers occurs in the ranking. Taking the median (*i.e.*, the 50% quantile) leads to the optimal, balanced, tradeoff studied in this paper.

- Second, $d_\tau(X_1; X_2)$ is equal to the minimum number of swaps of consecutive elements that are needed to transform the ranking induced by X_1 into the ranking induced by X_2 and vice versa. This interpretation is the cornerstone for geometric reasoning, as it makes the connection with the distance along the manifold (or path graph) of rankings.
- Third, $d_\tau(X_1; X_2)$ is linearly related to the rank correlation $\tau(X_1; X_2)$ as $d_\tau = \frac{1-\tau}{2}$.

Keeping these three interpretations in mind is helpful to derive Eqs. (13) to (15).

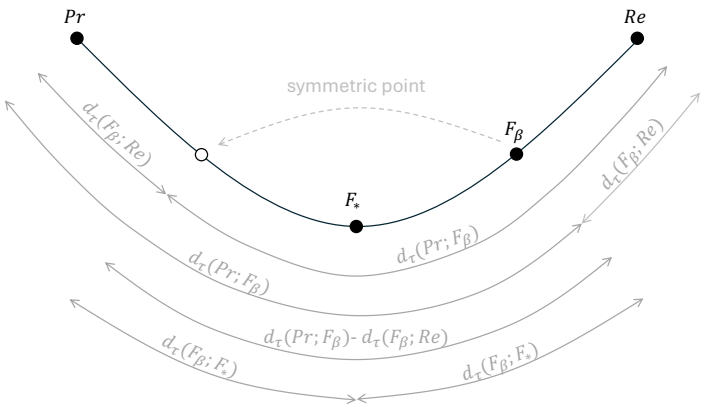
Let us consider Eq. (14), for example. We are interested in the probability $P(\text{✗})$. The first interpretation allows us to express it as $d_\tau(F_\beta; F_*)$. Then, it is possible to perform some geometric reasoning based on the second interpretation. As shown in Fig. A.2.3, $d_\tau(F_\beta; F_*) = {}^1\!/_2 \left| d_\tau(F_\beta; Re) - d_\tau(Pr; F_\beta) \right|$. Finally, using the third interpretation, we find that ${}^1\!/_2 \left| d_\tau(F_\beta; Re) - d_\tau(Pr; F_\beta) \right| = {}^1\!/_4 \left| \tau(Pr; F_\beta) - \tau(F_\beta; Re) \right|$.

A.2.4. On the CADA-RRE Example Used to Illustrate the Theory

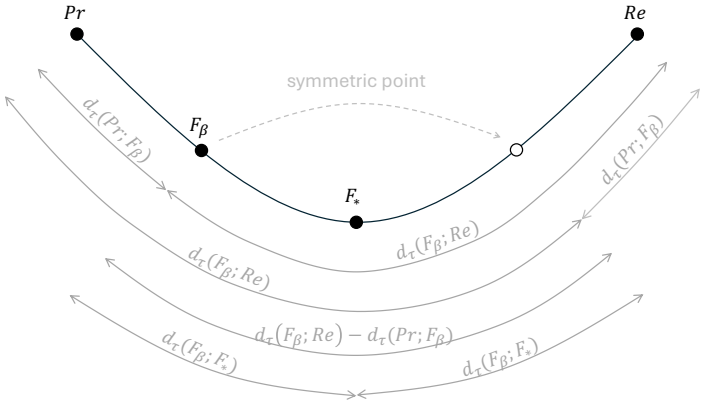
Among the various challenges hosted on *Grand Challenge*, we found that the third task of the *Cerebral Aneurysm Detection and Analysis* (CADA) challenge [17] is particularly interesting. This task is known as CADA-RRE for *Cerebral Aneurysm Rupture Risk Estimation* and is hosted at `https://cada-rre.grand-challenge.org/`. It consists in the automatic classification of known ruptured and unruptured aneurysms based on rotational X-ray angiographic images.

Even if the ranking chosen by the organizers is only based on F_2, the logs contain for each entry the values of five ranking scores [31]: the precision Pr, recall Re, accuracy A, F_1, and indeed F_2. This information allows one to recover the complete (normalized) confusion matrix with the probabilities of a true negative (PTN), false positive (PFP), false negative (PFN), and true positive (PTP). From these confusion matrices, we observed that the class priors are fixed and given by $\pi_- = {}^{19}\!/_{30}$ and $\pi_+ = {}^{11}\!/_{30}$. The values of PFP, PFN, and PTP can be used in Eq. (11), which gives the values used in Eq. (12) to compute an optimal β. The values of PFP and PFN can be used to derive $\mathbf{E}[PFP]$ and $\mathbf{E}[PFN]$, which in turn can be used to recommend a value for β with our heuristic Eq. (22).

At the time of writing these lines, the challenge leaderboard contains 29 entries. However, there are only 16 unique pairs of precision and recall values. This can be explained by the low number of cases on which the methods are tested. In our example based on CADA-RRE, we decided to rank the classifiers corresponding to these 16 unique pairs. This stands in contrast with



(a) Case in which the ranking induced by F_β is closer to the ranking induced by Re than to the ranking induced by Pr. We see that $d_\tau(F_\beta; F_*) = \frac{d_\tau(Pr;F_\beta) - d_\tau(F_\beta;Re)}{2}$.



(b) Case in which the ranking induced by F_β is closer to the ranking induced by Pr than to the ranking induced by Re. We see that $d_\tau(F_\beta; F_*) = \frac{d_\tau(F_\beta;Re) - d_\tau(Pr;F_\beta)}{2}$.

Figure A.2.3. Kendall distances on the manifold of rankings induced by the F_β scores. We see that $d_\tau(F_\beta; F_*) = \frac{\left|d_\tau(F_\beta;Re) - d_\tau(Pr;F_\beta)\right|}{2}$.

the official ranking of CADA-RRE as, for each participant, the organizers only considered the first three submissions with $F_2 \geq 0.1$. On the one hand, according to [17], the official ranking only contains three methods as only three teams submitted their solution: the 1st place is for [18] with $F_2 = 0.702$, the 2nd place is for [26] with $F_2 = 0.678$, and the 3rd place is for an unpublished method with $F_2 = 0.377$. On the other hand, the F_2 score is between 0.0 and 0.862 in the 16 entries that we rank. The results that we have obtained with the 16 performances to rank are provided in Figs. 3, A.2.2 and A.3.64.

What is particularly interesting with this example is that the organizers chose to rank the participants according to the F_2 score, as "the identification of aneurysms at risk is considered more important than the avoidance of false-positive risk classification". As shown in the plots of Fig. 3, the ranking induced by F_2 perfectly mimics the ranking induced by Re, totally ignoring the ranking induced by Pr. This is certainly not what is expected when one chooses to rank according to some F_β. As shown in Fig. A.2.2 and already discussed in Appendix A.2.2, $\beta = 2$ is the value to use if one wants to weight four times more Re than Pr from the *point of view of values*, but $\beta = 0.914$ is the value to use to weight four times more Re than Pr from the *point of view of ranks*.

A.3. Supplementary Material About Sec. 4.(Case Studies)

A.3.1. Principle of Analytical Computations for Kendall's rank correlations τ with Distributions of Performances

Let us consider two scores X_1, X_2. We denote the probability for them to order differently two (different) performances P_A and P_B drawn at random, independently, by $d_\tau(X_1; X_2)$. Kendall's rank correlation is given by $\tau(X_1; X_2) = 1 - 2d_\tau(X_1; X_2)$. When we consider a distribution of performances $\mathcal{P}$, instead of a finite set Π, we cannot compute $d_\tau(X_1; X_2)$ with Eq. (5). Instead, we can compute analytically the following probabilities:

$$\text{Proba}\left[X_1(P_A) < X_1(P_B), X_2(P_A) < X_2(P_B)\right] = \int_{P_A \in \mathbb{P}_{(\Omega,\Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega,\Sigma)}} \mathbf{1}_{X_1(P_A)<X_1(P_B), X_2(P_A)<X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) d P_A d P_B \,, \tag{37}$$

$$\text{Proba}\left[X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)\right] = \int_{P_A \in \mathbb{P}_{(\Omega,\Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega,\Sigma)}} \mathbf{1}_{X_1(P_A)<X_1(P_B), X_2(P_A)>X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) d P_A d P_B \,, \tag{38}$$

$$\text{Proba}\left[X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)\right] = \int_{P_A \in \mathbb{P}_{(\Omega,\Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega,\Sigma)}} \mathbf{1}_{X_1(P_A)>X_1(P_B), X_2(P_A)<X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) d P_A d P_B \,, \tag{39}$$

$$\text{Proba}\left[X_1(P_A) > X_1(P_B), X_2(P_A) > X_2(P_B)\right] = \int_{P_A \in \mathbb{P}_{(\Omega,\Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega,\Sigma)}} \mathbf{1}_{X_1(P_A)>X_1(P_B), X_2(P_A)>X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) d P_A d P_B \,, \tag{40}$$

where the symbol $\mathbf{1}$ denotes the indicator and $f_{\mathcal{P}}$ the probability density function related to the distribution $\mathcal{P}$.

Kendall's rank correlation τ is then given by

$$\tau(X_1; X_2) = 1 - 2\left(\text{Proba}\left[X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)\right] + \text{Proba}\left[X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)\right]\right) . \tag{41}$$

Note that P_A and P_B are exchangeable in these equalities since they are drawn independently from the same distribution. Thus, by symmetry, we have

$$\text{Proba}\left[X_1(P_A) < X_1(P_B), X_2(P_A) < X_2(P_B)\right] = \text{Proba}\left[X_1(P_A) > X_1(P_B), X_2(P_A) > X_2(P_B)\right] , \tag{42}$$

and

$$\text{Proba}\left[X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)\right] = \text{Proba}\left[X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)\right] . \tag{43}$$

Moreover, the four probabilities sum to one. Thus, computing one of these four probabilities suffices to be able to determine τ, and if we know τ, then we are able to recover the four probabilities:

$$\tau(X_1; X_2) = 4\,\text{Proba}\left[X_1(P_A) < X_1(P_B), X_2(P_A) < X_2(P_B)\right] - 1 \,, \tag{44}$$

$$\tau(X_1; X_2) = 1 - 4\,\text{Proba}\left[X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)\right] , \tag{45}$$

$$\tau(X_1; X_2) = 1 - 4\,\text{Proba}\left[X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)\right] , \tag{46}$$

$$\tau(X_1; X_2) = 4\,\text{Proba}\left[X_1(P_A) > X_1(P_B), X_2(P_A) > X_2(P_B)\right] - 1 \,. \tag{47}$$

The source codes for the analytical results provided hereafter are for *Wolfram 14.2* (*a.k.a. Mathematica*).

A.3.2. Detailed Results for the Uniform Distribution Over All Performances

To find the probabilities necessary to compute τ, we integrate over the tetrahedron. We start with a few initializations.

```
Tetra = Simplex[{{1, 0, 0, 0}, {0, 1, 0, 0}, {0, 0, 1, 0}, {0, 0, 0, 1}}]
Ppv[ptn_, pfp_, pfn_, ptp_] := ptp/(pfp + ptp)
Tpr[ptn_, pfp_, pfn_, ptp_] := ptp/(pfn + ptp)
Fone[ptn_, pfp_, pfn_, ptp_] := 2*ptp/(pfp + pfn + 2*ptp)

tot = Integrate[
    1,
    {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
    {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
]
```

We have

$$\tau(Pr; Re) = \frac{1}{3}, \tag{48}$$

as shown with the code:

```
pLtLt = Integrate[
    Boole[
        Ppv[ptn1, pfp1, pfn1, ptp1] < Ppv[ptn2, pfp2, pfn2, ptp2] &&
        Tpr[ptn1, pfp1, pfn1, ptp1] < Tpr[ptn2, pfp2, pfn2, ptp2]
    ],
    {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
    {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
] / tot
tau = 4 * pLtLt - 1
```

We have

$$\tau(Pr; F_1) = \frac{2}{3}, \tag{49}$$

as shown with the code:

```
pLtGt = Integrate[
    Boole[
        Ppv[ptn1, pfp1, pfn1, ptp1] < Ppv[ptn2, pfp2, pfn2, ptp2] &&
        Fone[ptn1, pfp1, pfn1, ptp1] > Fone[ptn2, pfp2, pfn2, ptp2]
    ],
    {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
    {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
] / tot
tau = 1 - 4 * pLtGt
```

And we have

$$\tau(F_1; Re) = \frac{2}{3}, \tag{50}$$

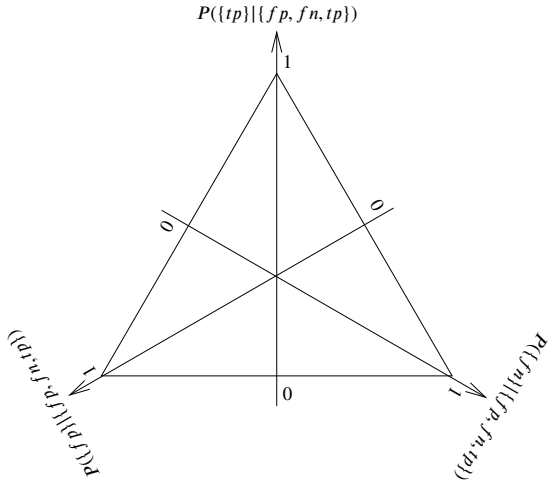
as shown with the code:

```
pLtGt = Integrate[
    Boole[
        Fone[ptn1, pfp1, pfn1, ptp1] < Fone[ptn2, pfp2, pfn2, ptp2] &&
        Tpr[ptn1, pfp1, pfn1, ptp1] > Tpr[ptn2, pfp2, pfn2, ptp2]
    ],
    {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
    {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
] / tot
tau = 1 - 4 * pLtGt
```

A.3.3. Detailed Results for the Uniform Distributions With Fixed Probability of True Negatives

The demonstration could be provided similarly as we did for the previous case study, taking a simplex in two dimensions instead of a simplex in three dimensions. However, we would like to show a geometrical reasoning instead.

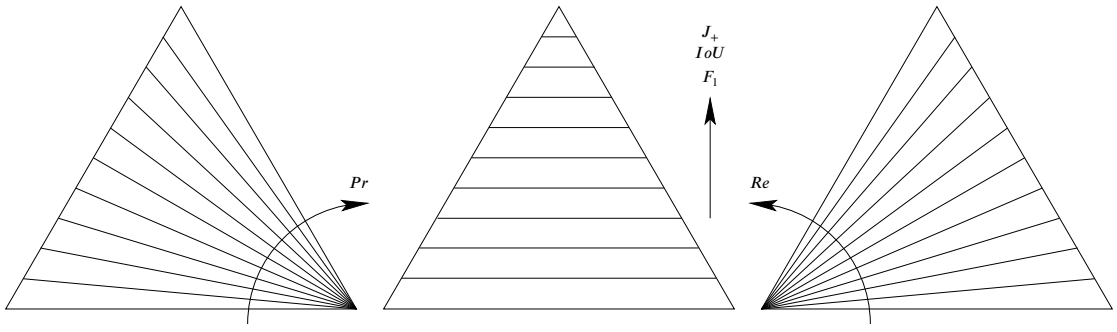
The equilateral triangles depicted in Fig. 4b form a "triangular space" in which the performances are assumed to be uniformly distributed. This "space" is parameterized as follows.



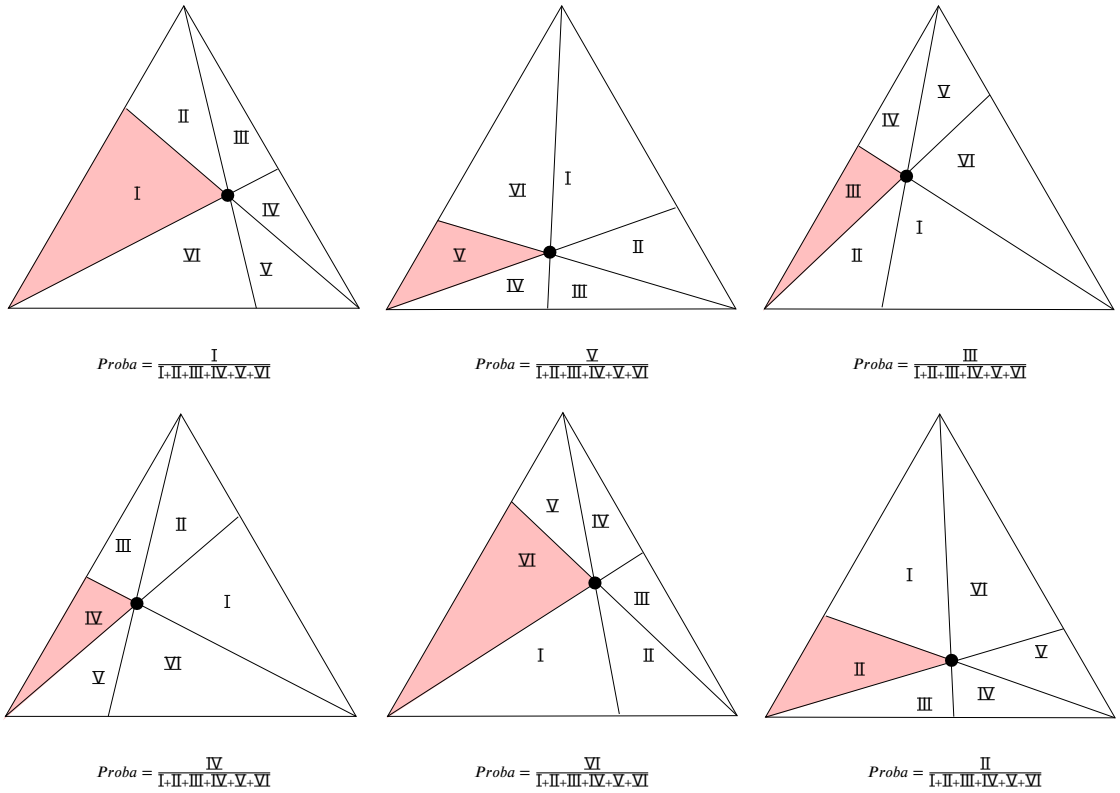
In fact, in the drawing shown here-above, the vertical axis corresponds to a score known as the *Intersection over Union* IoU or *Jaccard Coefficient* J_+. It is monotonically increasing with F_1:

$$F_1 = \frac{2IoU}{1 + IoU} \Rightarrow \frac{\partial F_1}{\partial IoU} > 0 \tag{51}$$

The isometrics of Pr form a pencil of lines whose vertex is located at the bottom-right corner. The isometrics of F_1 are horizontal lines. The isometrics of Re form a pencil of lines whose vertex is located at the bottom-left corner.



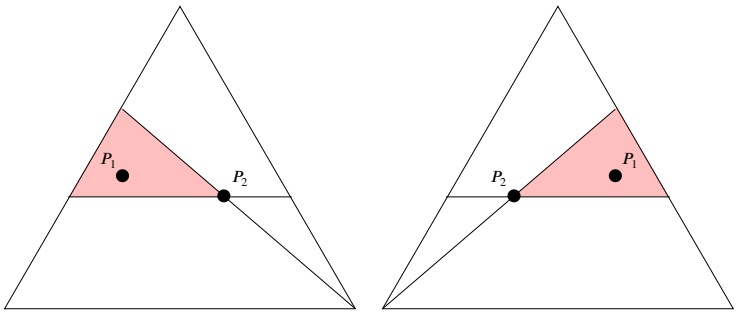
We now show that $\tau(Pr; Re) = 1/3$. In the following drawings, we consider a performance P_1 that can be anywhere in the "triangular space" and we choose one particular performance P_2, shown as a black point. The pink area represents the probability that $Pr(P_1) < Pr(P_2)$ and $Re(P_1) > Re(P_2)$. The six cases are symmetrical: one can obtain them by mirroring or rotating.



We see that, in average (that is, if we let P_2 be anywhere in the "triangular space"), the probability that $Pr(P_1) < Pr(P_2)$ and $Re(P_1) > Re(P_2)$ is $1/6$. Using Eq. (45), we find the announced result:

$$\tau(Pr; Re) = 1/3 \tag{52}$$

We now show that $\tau(Pr; F_1) = \tau(Re; F_1)$. Using Eq. (45), we know that it is the case if and only if the probability that $Pr(P_1) < Pr(P_2)$ and $F_1(P_1) > F_1(P_2)$ is equal to the probability that $Re(P_1) < Re(P_2)$ and $F_1(P_1) > F_1(P_2)$. These two probabilities are shown as pink areas, for some arbitrarily fixed P_2 (the black dot), in the following drawings.



We see that, by symmetry, the two probabilities are equal. If, furthermore, we do not fix anymore the performance P_2 but let it be anywhere in the "triangular space", the two probabilities remain equal. Using Eq. (45), we find the announced result:

$$\tau(Pr; F_1) = \tau(Re; F_1) \tag{53}$$

A.3.4. Detailed Results for the Uniform Distributions With Fixed Class Priors

To find the probabilities necessary to compute τ, we integrate over the whole ROC space. Let us denote by $(FPR, TPR) = (x, y)$ the coordinates of a performance in this space. We have

$$Pr(P_1) < Pr(P_2) \Leftrightarrow \frac{y_1}{x_1} < \frac{y_2}{x_2} \tag{54}$$

$$Re(P_1) < Re(P_2) \Leftrightarrow y_1 < y_2 \tag{55}$$

$$F_\beta(P_1) < F_\beta(P_2) \Leftrightarrow \frac{y_1}{x_1 + \ell} < \frac{y_2}{x_2 + \ell} \tag{56}$$

We start with a few initializations.

```
tot = Integrate[
    1,
    {x1, 0, 1}, {y1, 0, 1},
    {x2, 0, 1}, {y2, 0, 1}
]
```

We have

$$\tau(Pr; Re) = \frac{1}{2}, \tag{57}$$

as shown with the code (see Fig. A.3.4a):

```
pLtGt = Integrate[
        Boole[(y1 x2 < y2 x1) && (y1 > y2)],
        {x1, 0, 1}, {y1, 0, 1},
        {x2, 0, 1}, {y2, 0, 1}
] / tot
tau = 1 - 4 * pLtGt
```

We have

$$\tau(Pr; F_\beta) = 1 - \ell \left(\ell \log \left(\frac{\ell}{\ell + 1} \right) + 1 \right), \tag{58}$$

as shown with the code (see Fig. A.3.4b):

```
pLtGt = Integrate[
        Boole[(y1 x2 < y2 x1) && (y1 (x2 + 1) > y2 (x1 + 1))],
        {x1, 0, 1}, {y1, 0, 1},
        {x2, 0, 1}, {y2, 0, 1},
        Assumptions -> 1 > 0
] / tot
pLtGt = FullSimplify[pLtGt, Assumptions -> 1 > 0]
tau = 1 - 4 * pLtGt
```

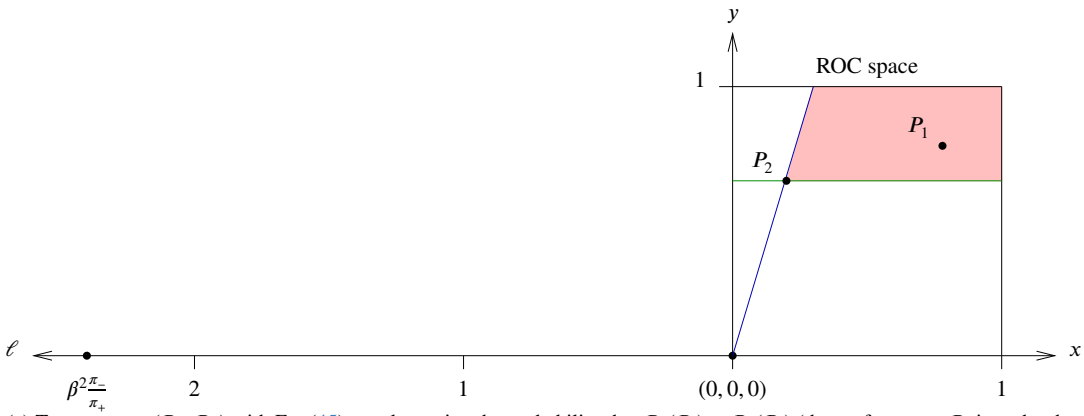
And we have

$$\tau(F_\beta; Re) = \ell^2 \left(-\log \left(\frac{1}{\ell} + 1 \right) \right) + \ell + \frac{1}{2}, \tag{59}$$

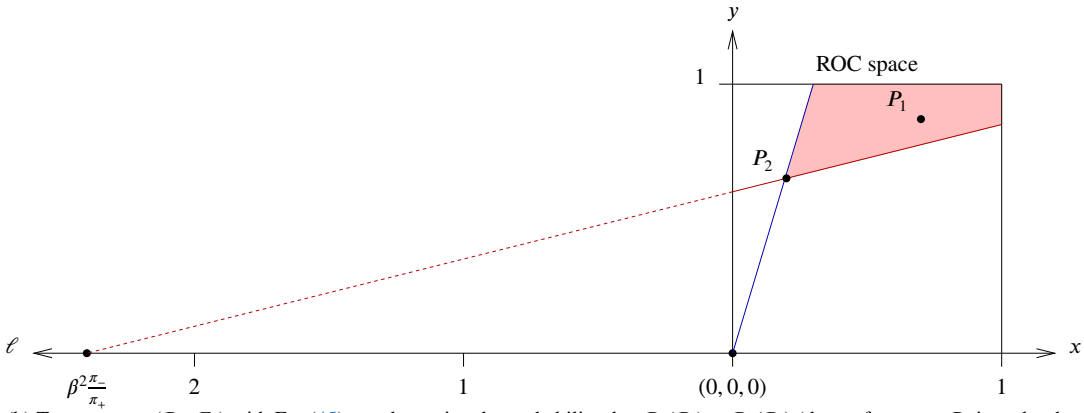
as shown with the code (see Fig. A.3.4c):

```
pLtGt = Integrate[
        Boole[(y1 (x2 + 1) < y2 (x1 + 1)) && (y1 > y2)],
        {x1, 0, 1}, {y1, 0, 1},
        {x2, 0, 1}, {y2, 0, 1},
        Assumptions -> 1 > 0
] / tot
tau = 1 - 4 * pLtGt
tau = FullSimplify[tau, Assumptions -> 1 > 0]
```

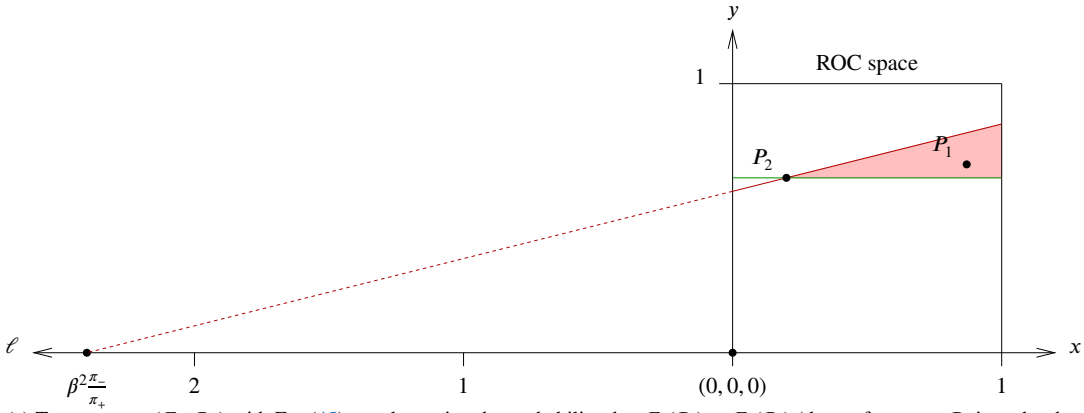
The plots are provided in Fig. A.3.5.



(a) To compute $\tau(Pr; Re)$ with Eq. (45), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2.

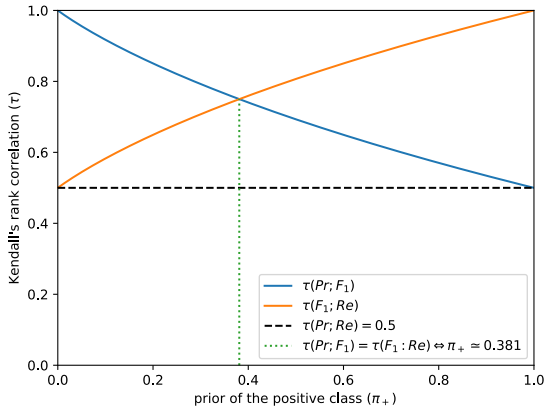


(b) To compute $\tau(Pr; F_\beta)$ with Eq. (45), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $F_\beta(P_1) > F_\beta(P_2)$ (the performance P_1 is above the red line), which is the pink area averaged for all positions of P_2.

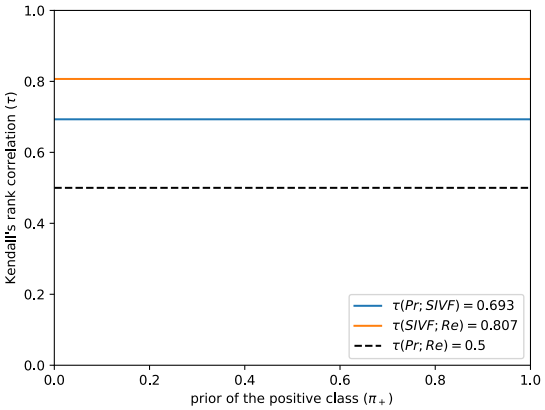


(c) To compute $\tau(F_\beta; Re)$ with Eq. (45), we determine the probability that $F_\beta(P_1) < F_\beta(P_2)$ (the performance P_1 is under the red line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2.

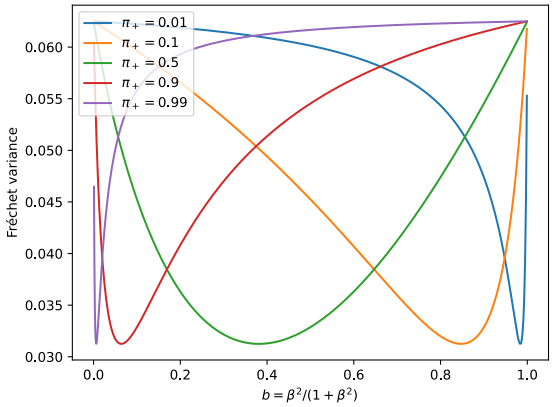
Figure A.3.4. Graphical representation, in and around the ROC space, of the principle that we use to derive the analytical expression for the optimal tradeoff between precision and recall for the purpose of ranking, in the case of uniform distributions with fixed class priors. Note that the ROC space is a linear mapping of the rectangles depicted in Fig. 4c.



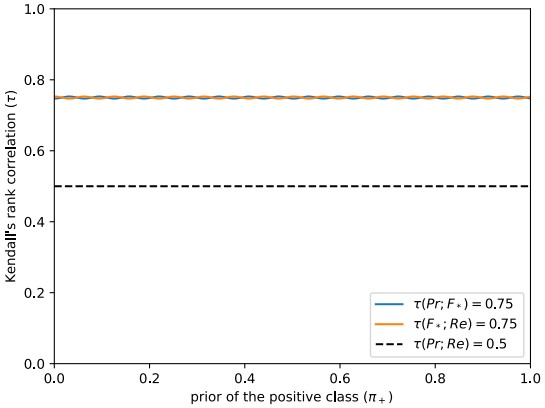
(a) F_1, the traditional (balanced) F-score, is not the optimal tradeoff: $\tau(Pr; F_1) \neq \tau(F_1; Re)$.



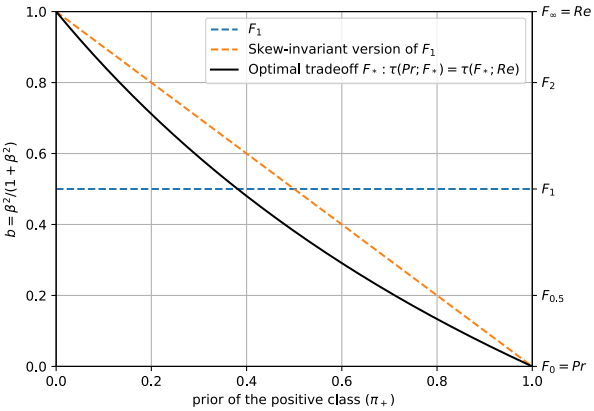
(b) $SIVF$, the skew-insensitive version of F_1 [10], is not the optimal tradeoff: $\tau(Pr; SIVF) \neq \tau(SIVF; Re)$.



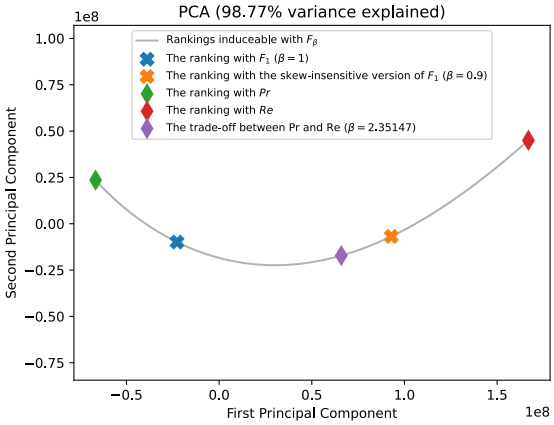
(c) Fréchet variance for various priors. It is defined at Eq. (8) and should be minimized to obtain the optimal tradeoff.



(d) F_* if the optimal tradeoff: $\tau(Pr; F_*) = \tau(F_*; Re)$.



(e) Adaptation of β w.r.t. class priors.



(f) PCA of the manifold (for $\pi_+ = 0.1$).

Figure A.3.5. Results for uniform distributions over the performances with fixed class priors, *i.e.* $\Pi_3(\pi_+)$.

A.3.5. Detailed Results for the Uniform Distributions With Fixed Class Priors, Above No-Skill

To find the probabilities necessary to compute τ, we integrate over half of the ROC space, the half above the raising diagonal. Let us denote by $(FPR, TPR) = (x, y)$ the coordinates of a performance in this space. We have

$$Pr(P_1) < Pr(P_2) \Leftrightarrow \frac{y_1}{x_1} < \frac{y_2}{x_2} \tag{60}$$

$$Re(P_1) < Re(P_2) \Leftrightarrow y_1 < y_2 \tag{61}$$

$$F_\beta(P_1) < F_\beta(P_2) \Leftrightarrow \frac{y_1}{x_1 + \ell} < \frac{y_2}{x_2 + \ell} \tag{62}$$

We start with a few initializations.

```
tot = Integrate[
    1,
    {x1, 0, 1}, {y1, x1, 1},
    {x2, 0, 1}, {y2, x2, 1}
]
```

We have

$$\tau(Pr; Re) = 0\,, \tag{63}$$

as shown with the code (see Fig. A.3.6a):

```
pLtGt = Integrate[
        Boole[(y1 x2 < y2 x1) && (y1 > y2)],
    {x1, 0, 1}, {y1, x1, 1},
    {x2, 0, 1}, {y2, x2, 1}
] / tot
tau = 1 - 4 * pLtGt
```

We have

$$\tau(Pr; F_\beta) = 1 - \frac{2}{3}\ell\left(-6\ell^2 - 6\left(\ell^2 - 1\right)\ell \log\left(\frac{\ell}{\ell + 1}\right) + 3\ell + 4\right)\,, \tag{64}$$

as shown with the code (see Fig. A.3.6b):

```
pLtGt = Integrate[
        Boole[(y1 x2 < y2 x1) && (y1 (x2 + 1) > y2 (x1 + 1))],
    {x1, 0, 1}, {y1, x1, 1},
    {x2, 0, 1}, {y2, x2, 1},
        Assumptions -> 1 > 0
] / tot
pLtGt = FullSimplify[pLtGt, Assumptions -> 1 > 0]
tau = 1 - 4 * pLtGt
```

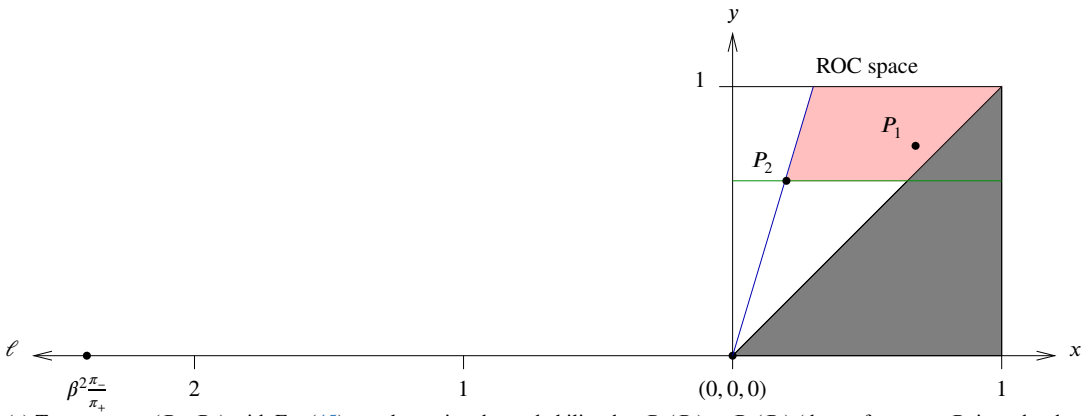
And we have

$$\tau(F_\beta; Re) = \frac{2}{3}\ell\left(-6\ell^2 + 6\left(\ell^2 - 1\right)\ell \log\left(\frac{1}{\ell} + 1\right) + 3\ell + 4\right)\,, \tag{65}$$

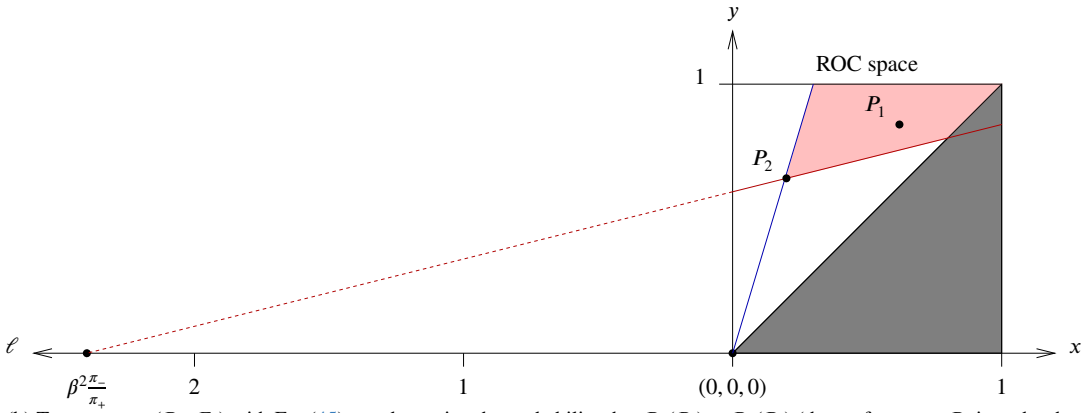
as shown with the code (see Fig. A.3.6c):

```
pLtGt = Integrate[
        Boole[(y1 (x2 + 1) < y2 (x1 + 1)) && (y1 > y2)],
    {x1, 0, 1}, {y1, x1, 1},
    {x2, 0, 1}, {y2, x2, 1},
        Assumptions -> 1 > 0
] / tot
tau = 1 - 4 * pLtGt
tau = FullSimplify[tau, Assumptions -> 1 > 0]
```

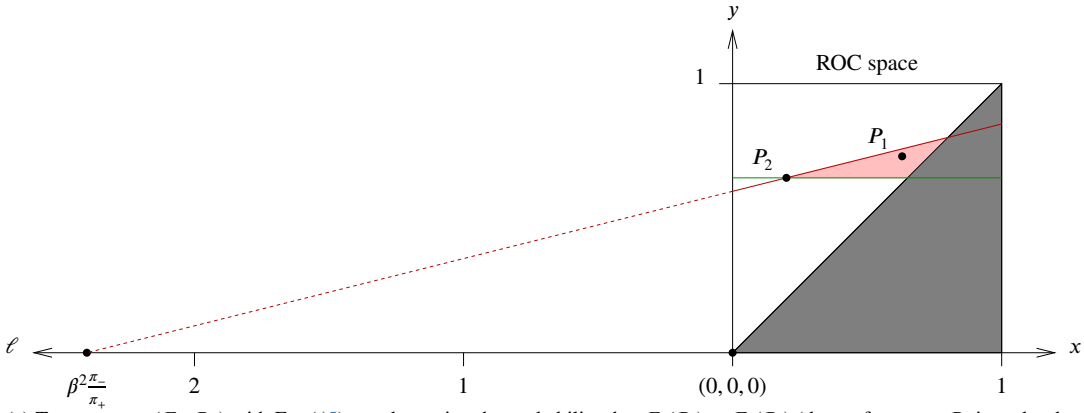
The plots are provided in Fig. A.3.7.



(a) To compute $\tau(Pr; Re)$ with Eq. (45), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2.

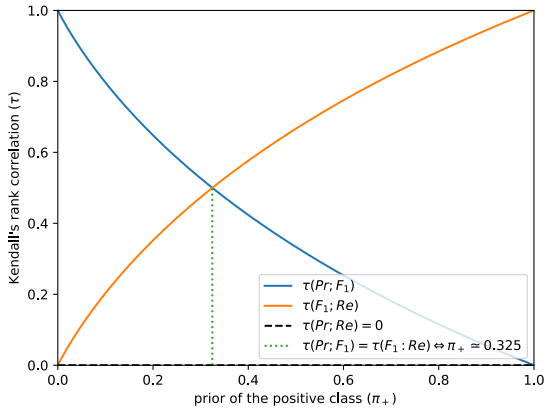


(b) To compute $\tau(Pr; F_\beta)$ with Eq. (45), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $F_\beta(P_1) > F_\beta(P_2)$ (the performance P_1 is above the red line), which is the pink area averaged for all positions of P_2.

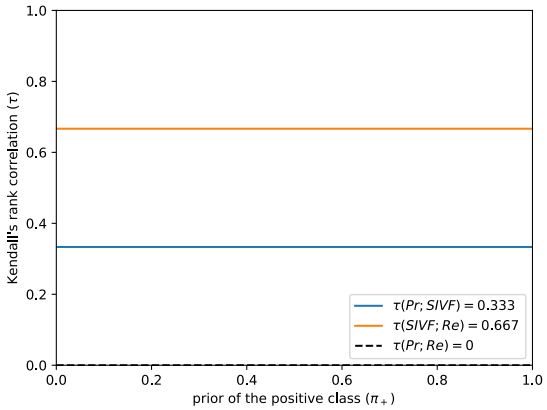


(c) To compute $\tau(F_\beta; Re)$ with Eq. (45), we determine the probability that $F_\beta(P_1) < F_\beta(P_2)$ (the performance P_1 is under the red line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2.

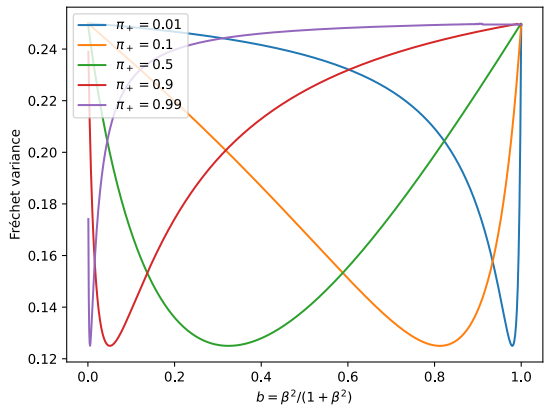
Figure A.3.6. Graphical representation, in and around the ROC space, of the principle that we use to derive the analytical expression for the optimal tradeoff between precision and recall for the purpose of ranking, in the case of uniform distributions with fixed class priors above no-skill. Note that the unshaded area of ROC space is a linear mapping of the triangles depicted in Fig. 4d.



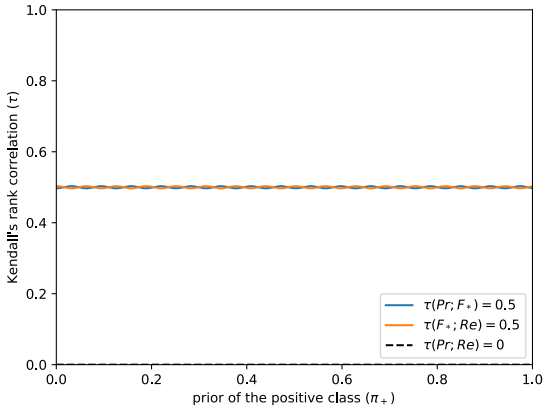
(a) F_1, the traditional (balanced) F-score, is not the optimal tradeoff: $\tau(Pr; F_1) \neq \tau(F_1; Re)$.



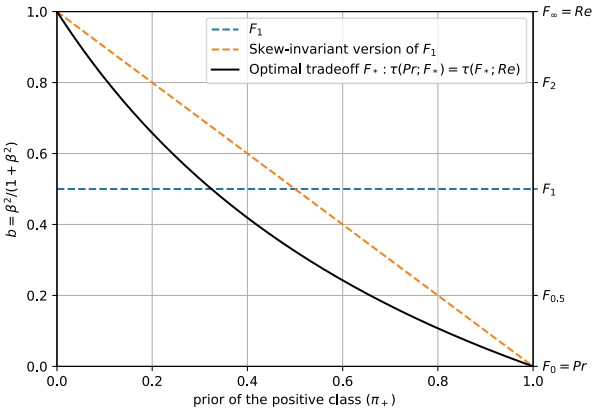
(b) $SIVF$, the skew-insensitive version of F_1 [10], is not the optimal tradeoff: $\tau(Pr; SIVF) \neq \tau(SIVF; Re)$.



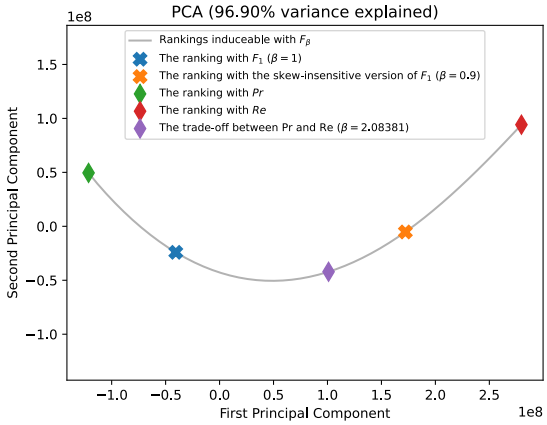
(c) Fréchet variance for various priors. It is defined at Eq. (8) and should be minimized to obtain the optimal tradeoff.



(d) F_* if the optimal tradeoff: $\tau(Pr; F_*) = \tau(F_*; Re)$.



(e) Adaptation of β w.r.t. class priors.



(f) PCA of the manifold (for $\pi_+ = 0.1$).

Figure A.3.7. Results for uniform distributions over the performances with fixed class priors, *i.e.* $\Pi_4(\pi_+)$.

A.3.6. Detailed Results for the Uniform Distributions With Fixed Class Priors, Close to Oracle

To find the probabilities necessary to compute τ, we integrate over a rectangle in the ROC space. Let us denote by $(FPR, TPR) = (x, y)$ the coordinates of a performance in this space. We have

$$Pr(P_1) < Pr(P_2) \Leftrightarrow \frac{y_1}{x_1} < \frac{y_2}{x_2} \tag{66}$$

$$Re(P_1) < Re(P_2) \Leftrightarrow y_1 < y_2 \tag{67}$$

$$F_\beta(P_1) < F_\beta(P_2) \Leftrightarrow \frac{y_1}{x_1 + \ell} < \frac{y_2}{x_2 + \ell} \qquad \text{with} \qquad \ell = \beta^2 \frac{\pi_+}{\pi_-} = \frac{b}{1-b} \frac{\pi_+}{1-\pi_+} \tag{68}$$

Unlike the previous cases, here we can no longer group π_+ and β together to form ℓ, as π_+ appears alone in the integration bounds of x and y. More precisely, we cannot anymore minimize Fréchet variance as a function of just ℓ. The optimal ℓ is not a constant anymore: it is now a function of π_+. This explains why the look of the curve for the adaptation differs from what we had before.

We start with a few initializations.

```
tot = Integrate[
    1,
    {x1, 0, p}, {y1, p, 1},
    {x2, 0, p}, {y2, p, 1},
        Assumptions -> 0 < p < 1
]
```

We have

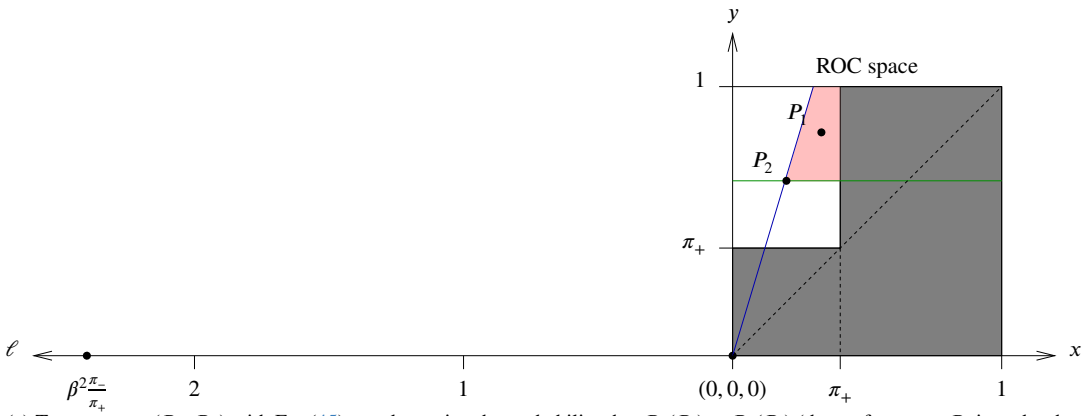
$$\tau(Pr; Re) = 1 - \frac{-\pi_+^4 + 2\pi_+^4 \log(\pi_+) + \pi_+^2}{2(1-\pi_+)^2 \pi_+^2}, \tag{69}$$

as shown with the code (see Fig. A.3.8a):

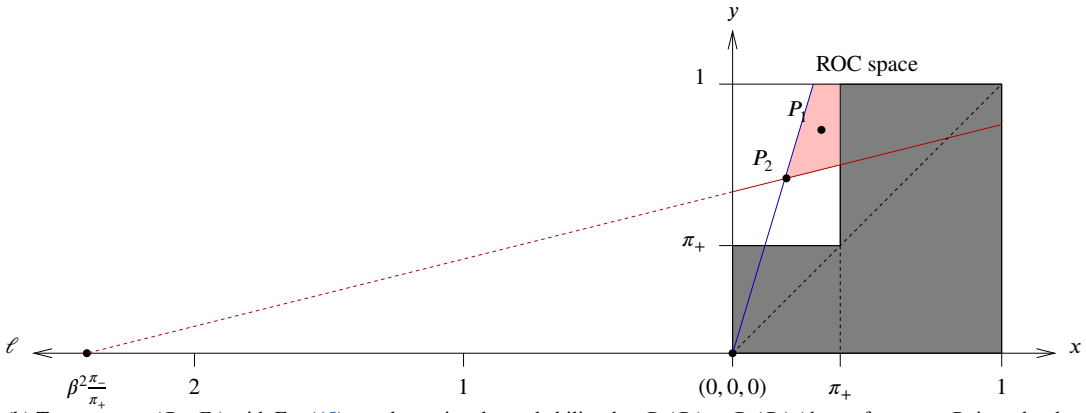
```
pLtGt = Integrate[
        Boole[(y1 x2 < y2 x1) && (y1 > y2)],
    {x1, 0, p}, {y1, p, 1},
    {x2, 0, p}, {y2, p, 1},
        Assumptions -> 0 < p < 1
] / tot
tau = 1 - 4 * pLtGt
```

With Wolfram, we were unable to obtain the analytical expression of $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ for this family of distributions. Wolfram was unable to provide you with the analytical expression in a reasonable time. We therefore report the results obtained using the Monte-Carlo technique.

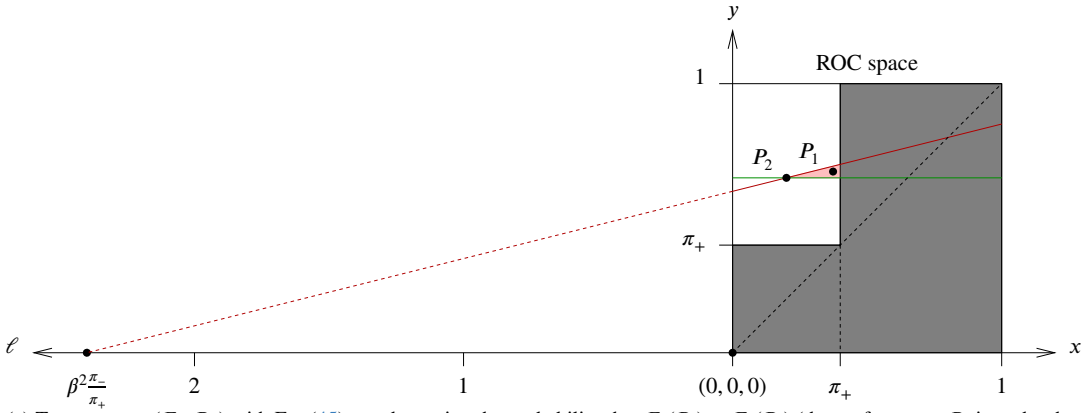
The plots are provided in Fig. A.3.9.



(a) To compute $\tau(Pr; Re)$ with Eq. (45), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2.

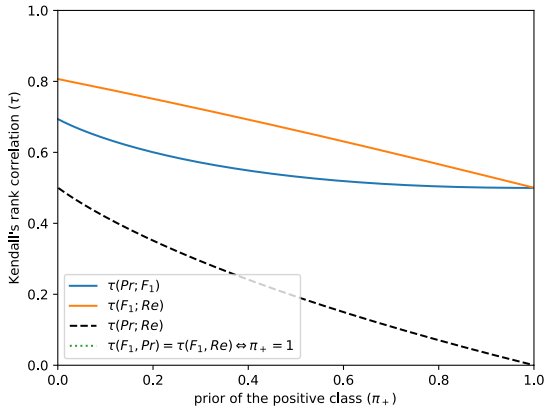


(b) To compute $\tau(Pr; F_\beta)$ with Eq. (45), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $F_\beta(P_1) > F_\beta(P_2)$ (the performance P_1 is above the red line), which is the pink area averaged for all positions of P_2.

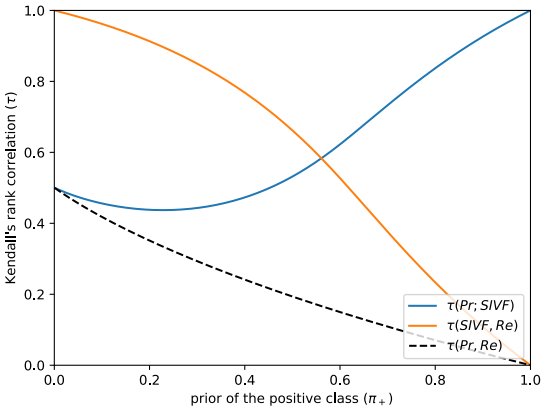


(c) To compute $\tau(F_\beta; Re)$ with Eq. (45), we determine the probability that $F_\beta(P_1) < F_\beta(P_2)$ (the performance P_1 is under the red line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2.

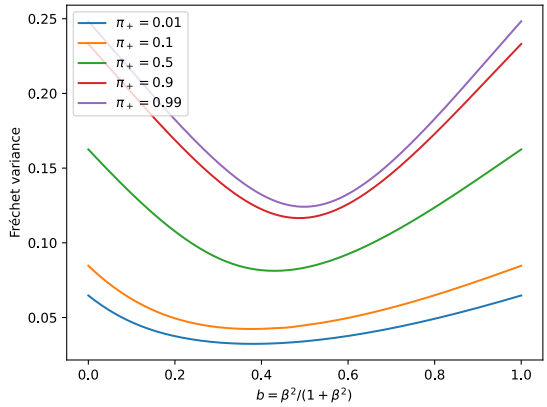
Figure A.3.8. Graphical representation, in and around the ROC space, of the principle that we use to derive the analytical expression for the optimal tradeoff between precision and recall for the purpose of ranking, in the case of uniform distributions with fixed class priors close to oracle. Note that the unshaded area of ROC space is a linear mapping of the squares depicted in Fig. 4e.



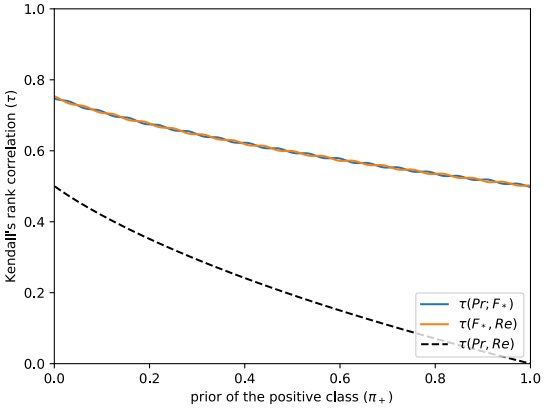
(a) F_1, the traditional (balanced) F-score, is not the optimal tradeoff: $\tau(Pr; F_1) \neq \tau(F_1; Re)$.



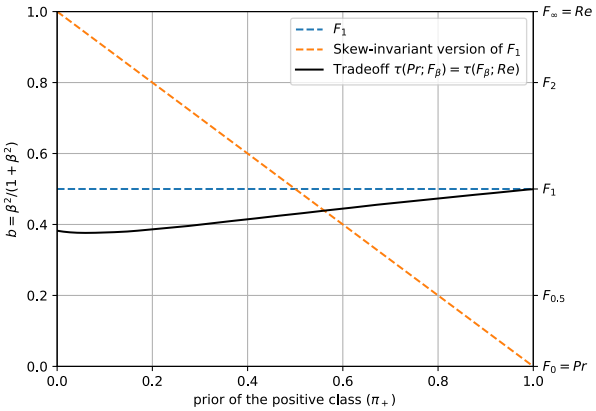
(b) $SIVF$, the skew-insensitive version of F_1 [10], is not the optimal tradeoff: $\tau(Pr; SIVF) \neq \tau(SIVF; Re)$.



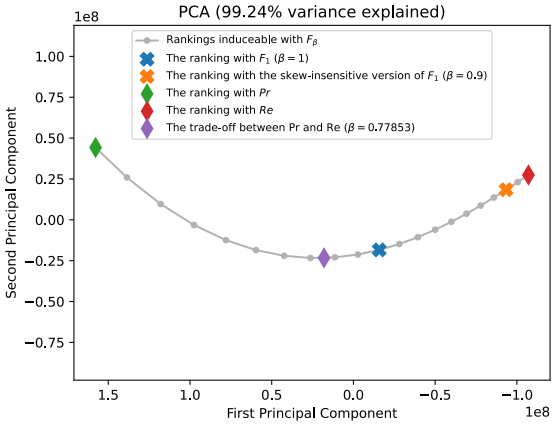
(c) Fréchet variance for various priors. It is defined at Eq. (8) and should be minimized to obtain the optimal tradeoff.



(d) F_* if the optimal tradeoff: $\tau(Pr; F_*) = \tau(F_*; Re)$.



(e) Adaptation of β w.r.t. class priors.



(f) PCA of the manifold (for $\pi_+ = 0.1$).

Figure A.3.9. Results for uniform distributions over the performances with fixed class priors, *i.e.* $\Pi_5(\pi_+)$.

A.3.7. All Results for Some Real Sets of Performances

We provide here the detailed results for each video of *CDnet 2014*.

Category "Baseline"

- The results for video "pedestrians" (ranking of 57 methods; $\pi_+ = 0.0098$) can be found at Fig. A.3.10.
- The results for video "PETS2006" (ranking of 57 methods; $\pi_+ = 0.0130$) can be found at Fig. A.3.11.
- The results for video "office" (ranking of 57 methods; $\pi_+ = 0.0690$) can be found at Fig. A.3.12.
- The results for video "highway" (ranking of 57 methods; $\pi_+ = 0.0593$) can be found at Fig. A.3.13.

Category "Dynamic Background"

- The results for video "overpass" (ranking of 57 methods; $\pi_+ = 0.0134$) can be found at Fig. A.3.14.
- The results for video "canoe" (ranking of 57 methods; $\pi_+ = 0.0354$) can be found at Fig. A.3.15.
- The results for video "fountain01" (ranking of 57 methods; $\pi_+ = 0.0008$) can be found at Fig. A.3.16.
- The results for video "fountain02" (ranking of 57 methods; $\pi_+ = 0.0022$) can be found at Fig. A.3.17.
- The results for video "fall" (ranking of 57 methods; $\pi_+ = 0.0177$) can be found at Fig. A.3.18.
- The results for video "boats" (ranking of 57 methods; $\pi_+ = 0.0063$) can be found at Fig. A.3.19.

Category "Camera Jitter"

- The results for video "boulevard" (ranking of 57 methods; $\pi_+ = 0.0469$) can be found at Fig. A.3.20.
- The results for video "sidewalk" (ranking of 57 methods; $\pi_+ = 0.0261$) can be found at Fig. A.3.21.
- The results for video "badminton" (ranking of 57 methods; $\pi_+ = 0.0343$) can be found at Fig. A.3.22.
- The results for video "traffic" (ranking of 57 methods; $\pi_+ = 0.0623$) can be found at Fig. A.3.23.

Category "Intermittent Object Motion"

- The results for video "abandonedBox" (ranking of 57 methods; $\pi_+ = 0.0481$) can be found at Fig. A.3.24.
- The results for video "winterDriveway" (ranking of 57 methods; $\pi_+ = 0.0075$) can be found at Fig. A.3.25.
- The results for video "sofa" (ranking of 57 methods; $\pi_+ = 0.0437$) can be found at Fig. A.3.26.
- The results for video "tramstop" (ranking of 57 methods; $\pi_+ = 0.1795$) can be found at Fig. A.3.27.
- The results for video "parking" (ranking of 57 methods; $\pi_+ = 0.0773$) can be found at Fig. A.3.28.
- The results for video "streetLight" (ranking of 57 methods; $\pi_+ = 0.0485$) can be found at Fig. A.3.29.

Category "Shadow"

- The results for video "copyMachine" (ranking of 57 methods; $\pi_+ = 0.0693$) can be found at Fig. A.3.30.
- The results for video "bungalows" (ranking of 57 methods; $\pi_+ = 0.0600$) can be found at Fig. A.3.31.
- The results for video "busStation" (ranking of 57 methods; $\pi_+ = 0.0369$) can be found at Fig. A.3.32.
- The results for video "peopleInShade" (ranking of 57 methods; $\pi_+ = 0.0564$) can be found at Fig. A.3.33.
- The results for video "backdoor" (ranking of 57 methods; $\pi_+ = 0.0199$) can be found at Fig. A.3.34.
- The results for video "cubicle" (ranking of 57 methods; $\pi_+ = 0.0196$) can be found at Fig. A.3.35.

Category "Thermal"

- The results for video "lakeSide" (ranking of 57 methods; $\pi_+ = 0.0192$) can be found at Fig. A.3.36.
- The results for video "diningRoom" (ranking of 57 methods; $\pi_+ = 0.0859$) can be found at Fig. A.3.37.
- The results for video "park" (ranking of 57 methods; $\pi_+ = 0.0203$) can be found at Fig. A.3.38.
- The results for video "corridor" (ranking of 57 methods; $\pi_+ = 0.0331$) can be found at Fig. A.3.39.
- The results for video "library" (ranking of 57 methods; $\pi_+ = 0.1928$) can be found at Fig. A.3.40.

Category "Bad Weather"

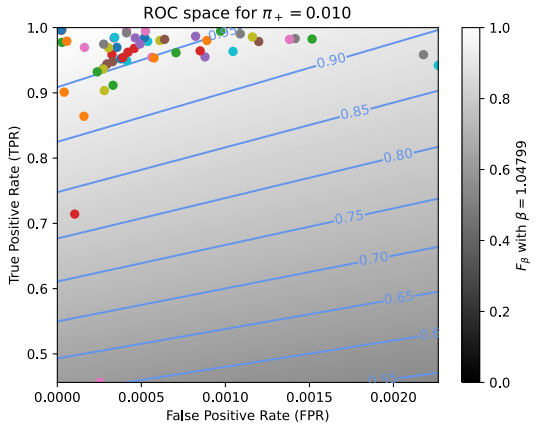
- The results for video "skating" (ranking of 57 methods; $\pi_+ = 0.0397$) can be found at Fig. A.3.41.
- The results for video "wetSnow" (ranking of 57 methods; $\pi_+ = 0.0150$) can be found at Fig. A.3.42.
- The results for video "snowFall" (ranking of 57 methods; $\pi_+ = 0.0105$) can be found at Fig. A.3.43.
- The results for video "blizzard" (ranking of 57 methods; $\pi_+ = 0.0115$) can be found at Fig. A.3.44.

Category "Low Framerate"
- The results for video "tunnelExit_0_35fps" (ranking of 57 methods; $\pi_+ = 0.0195$) can be found at Fig. A.3.45.
- The results for video "port_0_17fps" (ranking of 57 methods; $\pi_+ = 0.0002$) can be found at Fig. A.3.46.
- The results for video "tramCrossroad_1fps" (ranking of 57 methods; $\pi_+ = 0.0288$) can be found at Fig. A.3.47.
- The results for video "turnpike_0_5fps" (ranking of 57 methods; $\pi_+ = 0.0581$) can be found at Fig. A.3.48.

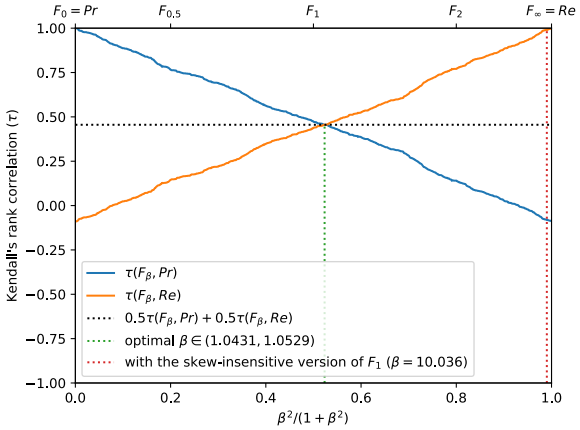
Category "Night Videos"
- The results for video "tramStation" (ranking of 58 methods; $\pi_+ = 0.0339$) can be found at Fig. A.3.49.
- The results for video "busyBoulvard" (ranking of 58 methods; $\pi_+ = 0.0847$) can be found at Fig. A.3.50.
- The results for video "streetCornerAtNight" (ranking of 58 methods; $\pi_+ = 0.0062$) can be found at Fig. A.3.51.
- The results for video "fluidHighway" (ranking of 58 methods; $\pi_+ = 0.0175$) can be found at Fig. A.3.52.
- The results for video "bridgeEntry" (ranking of 58 methods; $\pi_+ = 0.0200$) can be found at Fig. A.3.53.
- The results for video "winterStreet" (ranking of 58 methods; $\pi_+ = 0.0689$) can be found at Fig. A.3.54.

Category "PTZ"
- The results for video "twoPositionPTZCam" (ranking of 56 methods; $\pi_+ = 0.0117$) can be found at Fig. A.3.55.
- The results for video "zoomInZoomOut" (ranking of 56 methods; $\pi_+ = 0.0019$) can be found at Fig. A.3.56.
- The results for video "continuousPan" (ranking of 56 methods; $\pi_+ = 0.0056$) can be found at Fig. A.3.57.
- The results for video "intermittentPan" (ranking of 56 methods; $\pi_+ = 0.0094$) can be found at Fig. A.3.58.

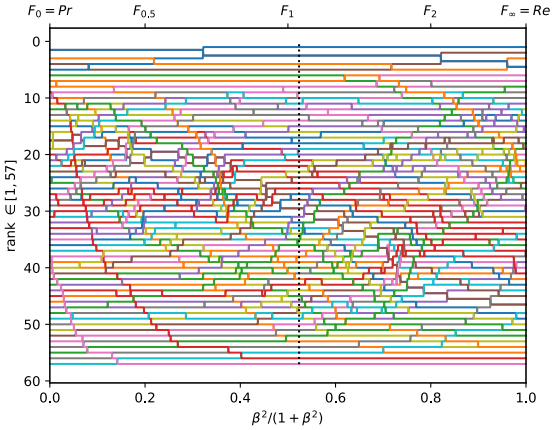
Category "Turbulence"
- The results for video "turbulence2" (ranking of 57 methods; $\pi_+ = 0.0008$) can be found at Fig. A.3.59.
- The results for video "turbulence3" (ranking of 57 methods; $\pi_+ = 0.0121$) can be found at Fig. A.3.60.
- The results for video "turbulence0" (ranking of 57 methods; $\pi_+ = 0.0015$) can be found at Fig. A.3.61.
- The results for video "turbulence1" (ranking of 57 methods; $\pi_+ = 0.0026$) can be found at Fig. A.3.62.



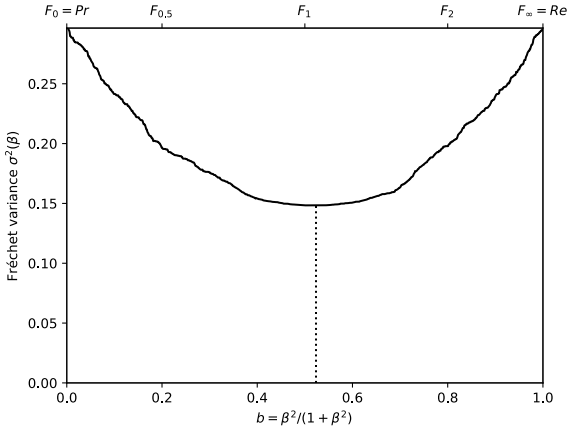
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



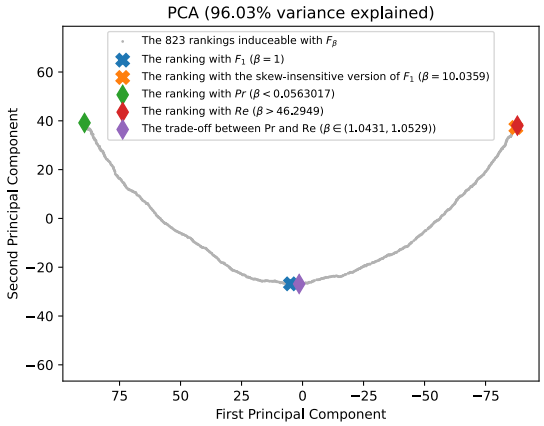
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



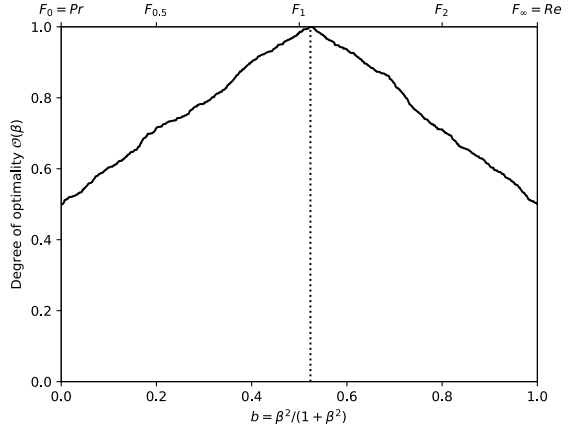
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

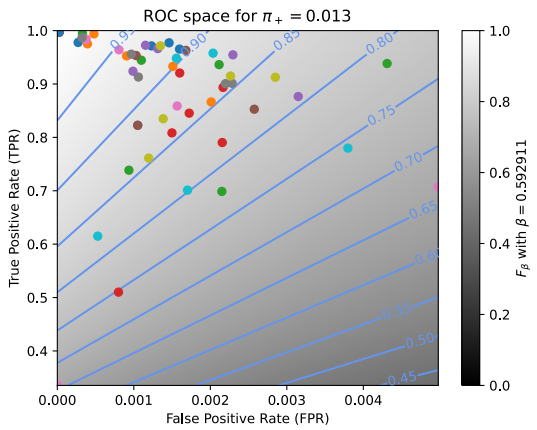


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

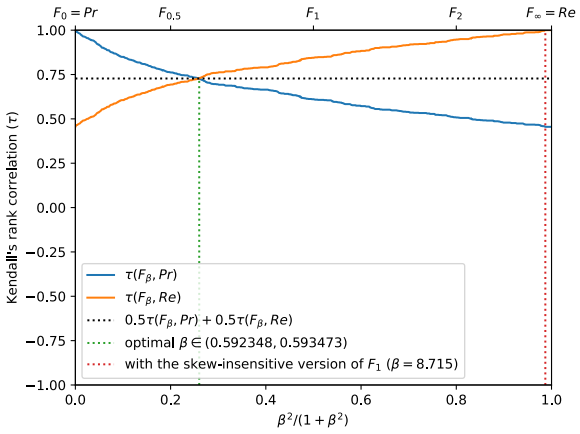


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

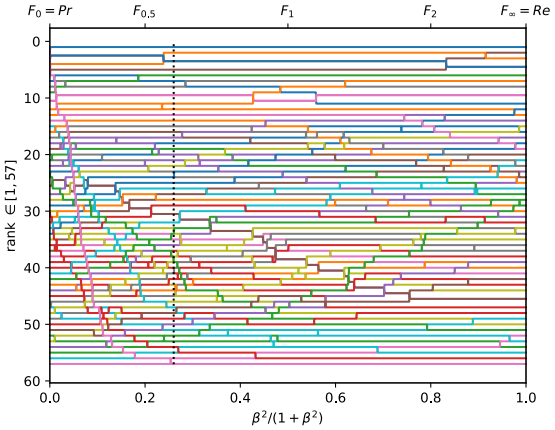
Figure A.3.10. Ranking of 57 BGS methods evaluated on the video "pedestrians" ($\pi_+ = 0.0098$).



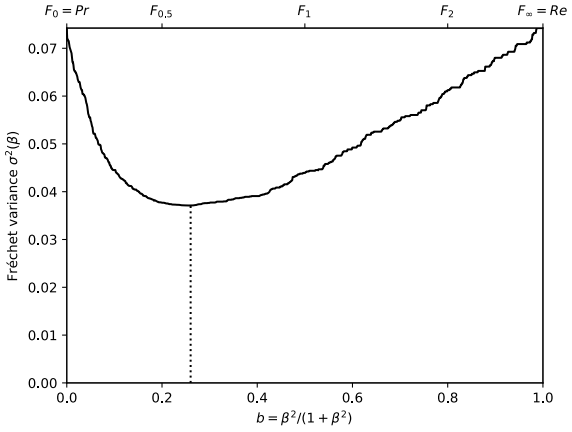
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



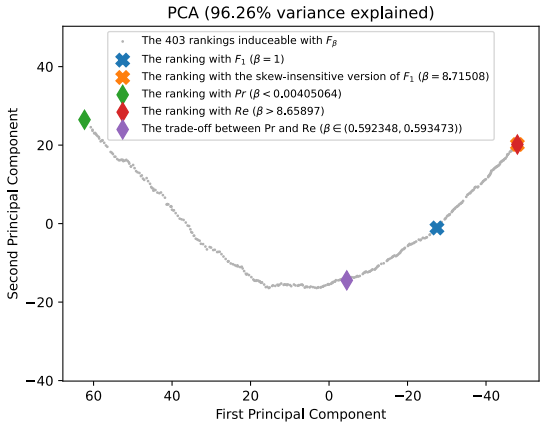
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



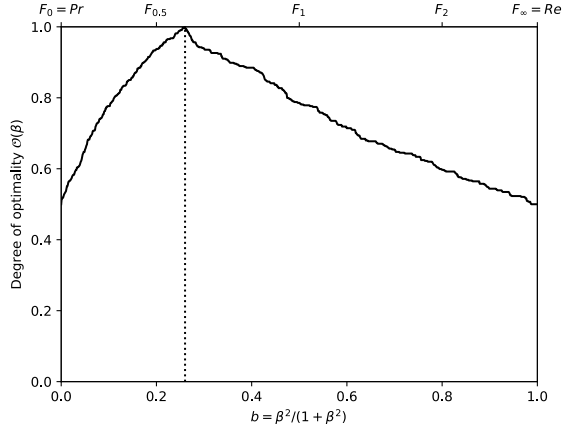
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

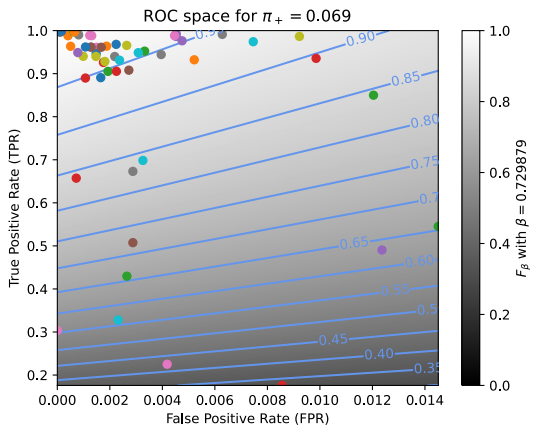


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

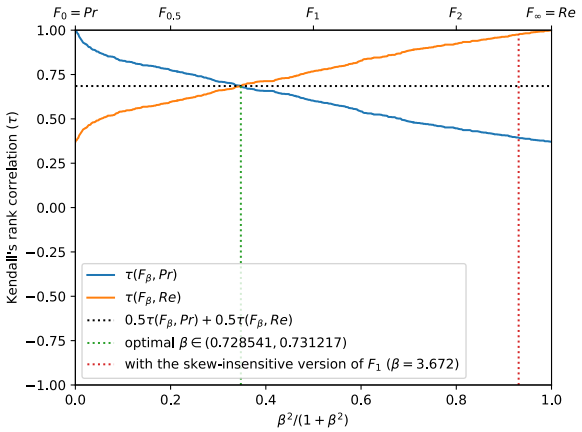


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

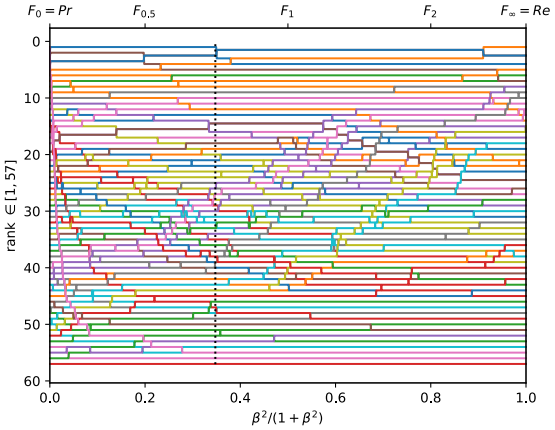
Figure A.3.11. Ranking of 57 BGS methods evaluated on the video "PETS2006" ($\pi_+ = 0.0130$).



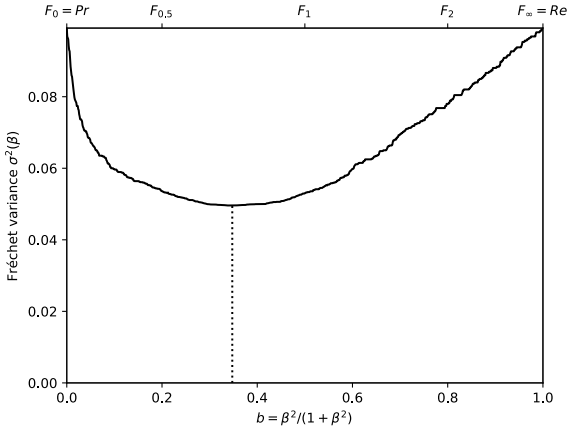
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



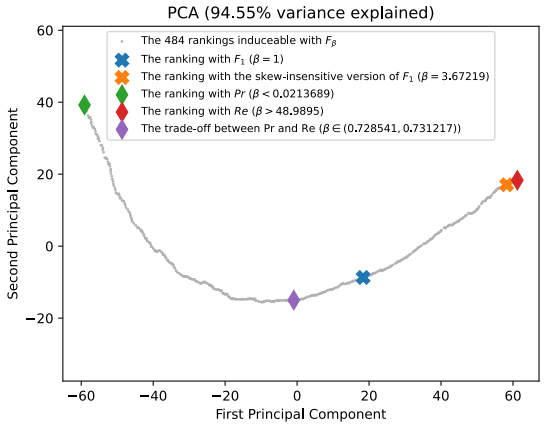
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



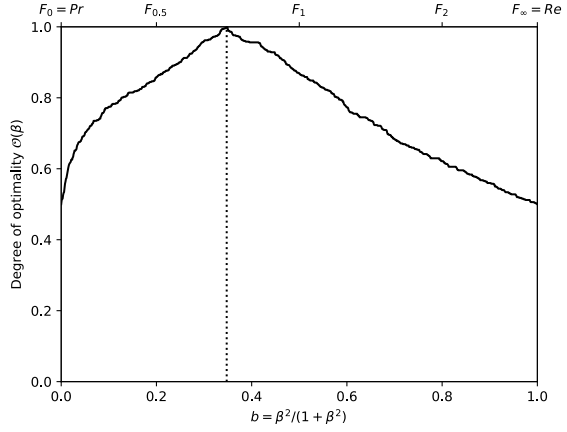
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

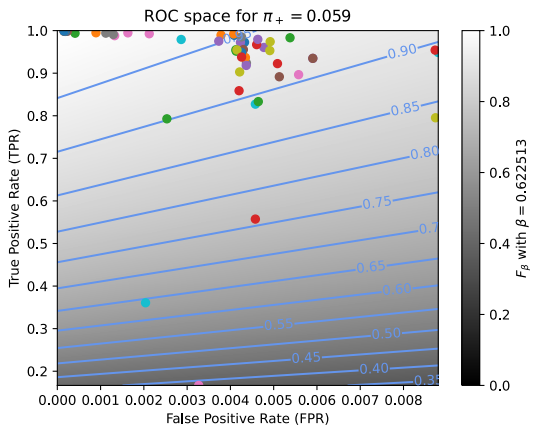


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

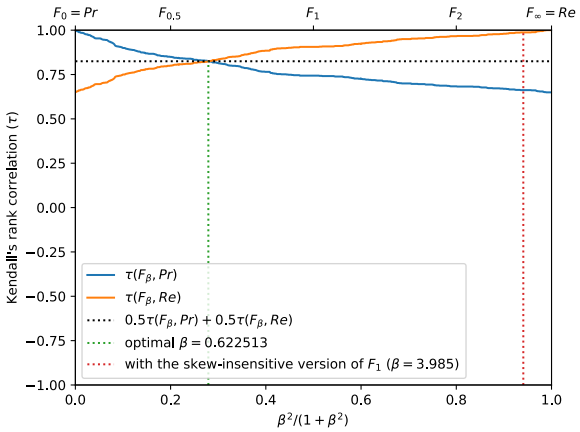


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

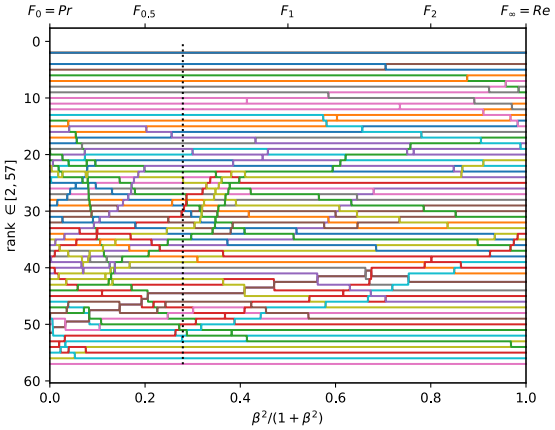
Figure A.3.12. Ranking of 57 BGS methods evaluated on the video "office" ($\pi_+ = 0.0690$).



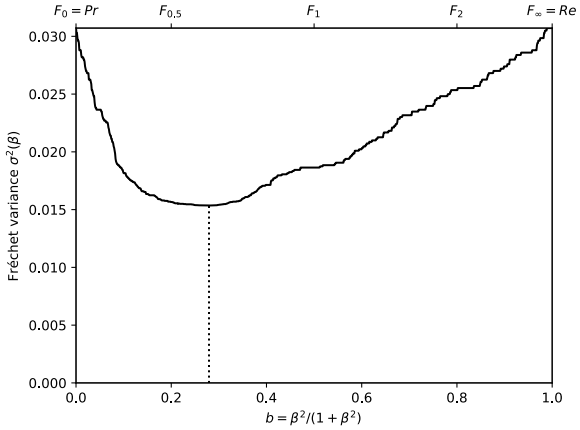
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



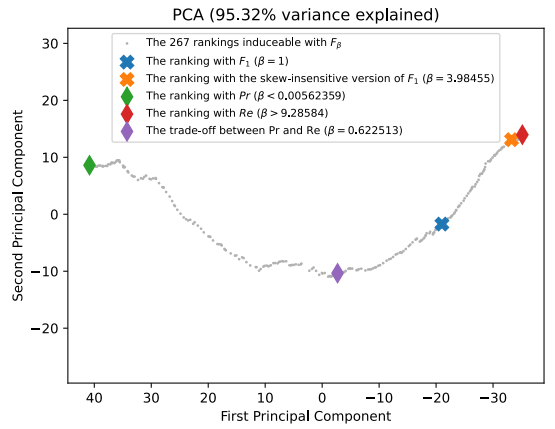
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



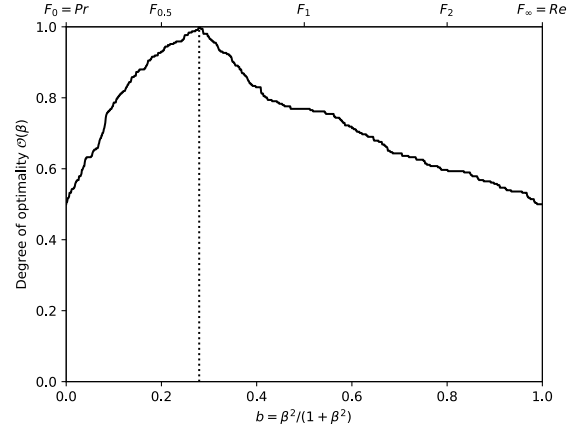
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

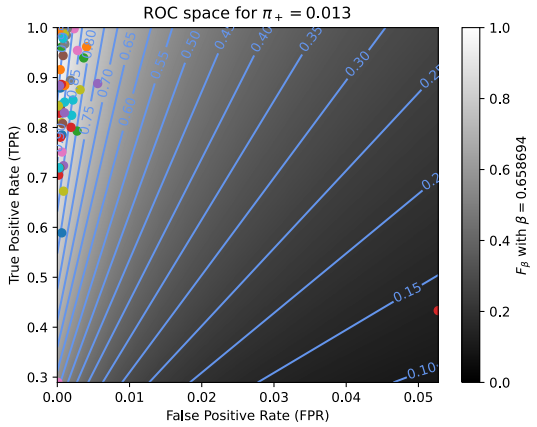


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

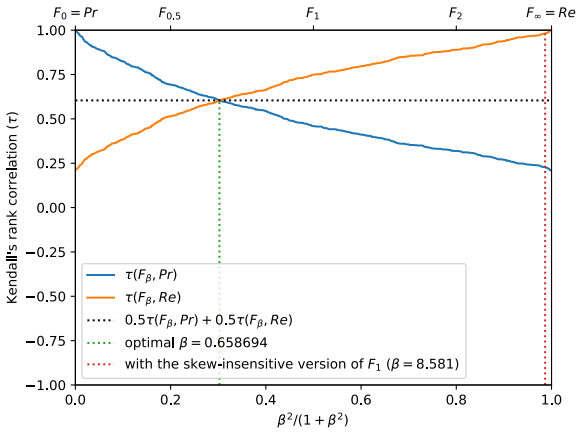


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

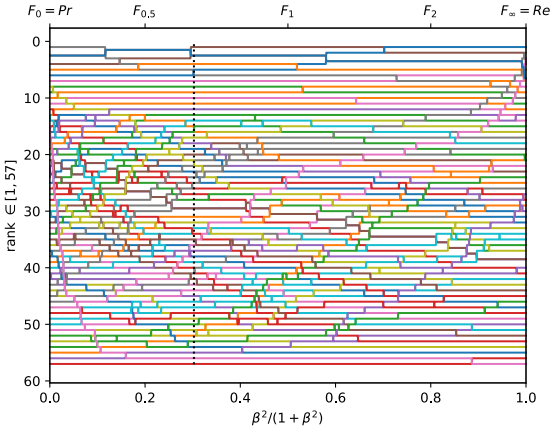
Figure A.3.13. Ranking of 57 BGS methods evaluated on the video "highway" ($\pi_+ = 0.0593$).



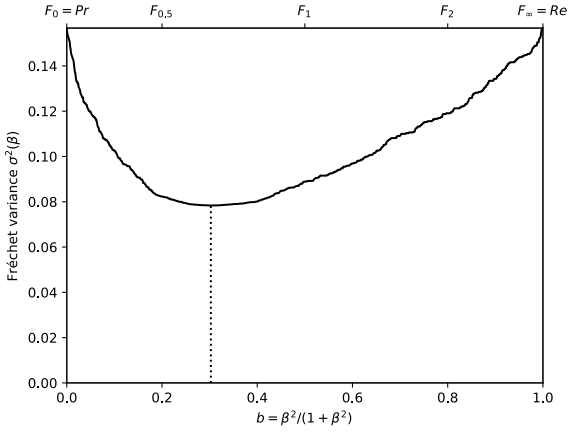
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



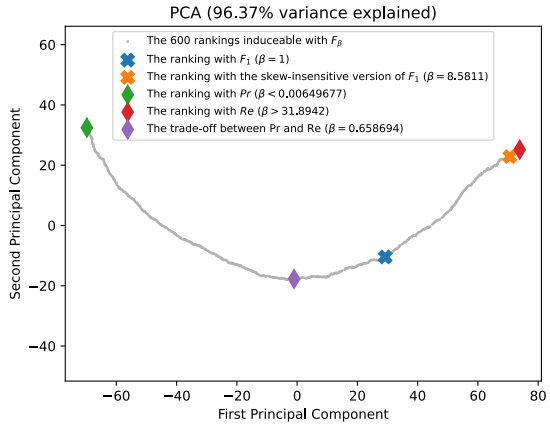
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



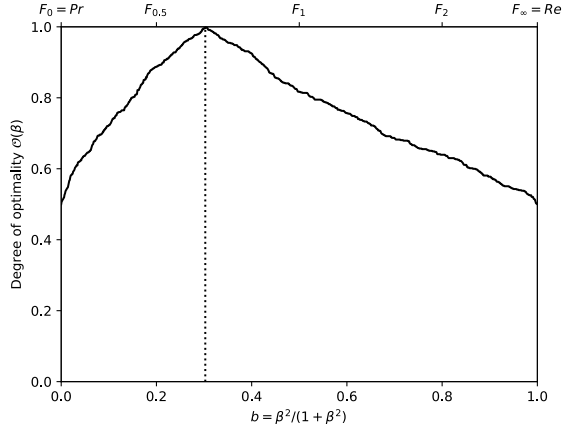
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

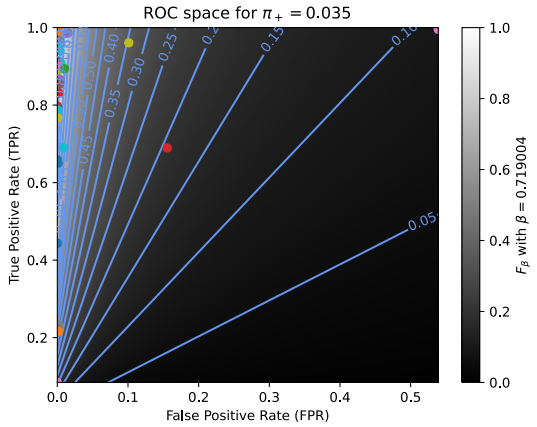


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

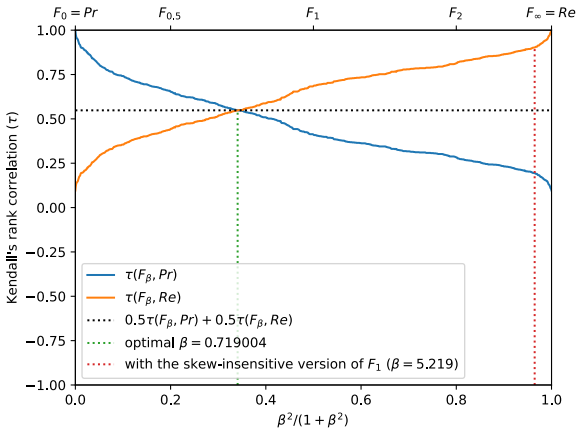


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

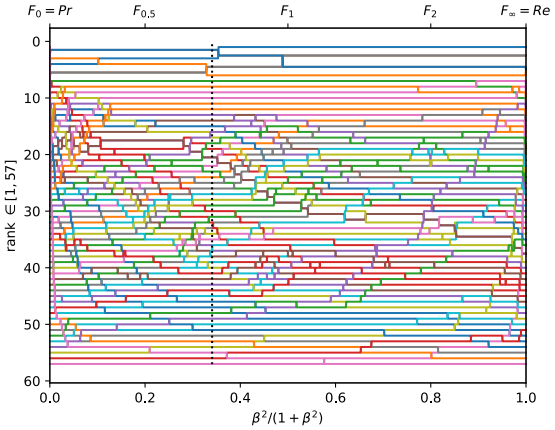
Figure A.3.14. Ranking of 57 BGS methods evaluated on the video "overpass" ($\pi_+ = 0.0134$).



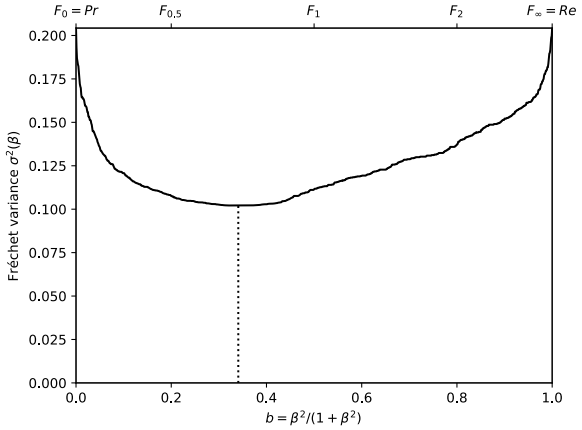
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



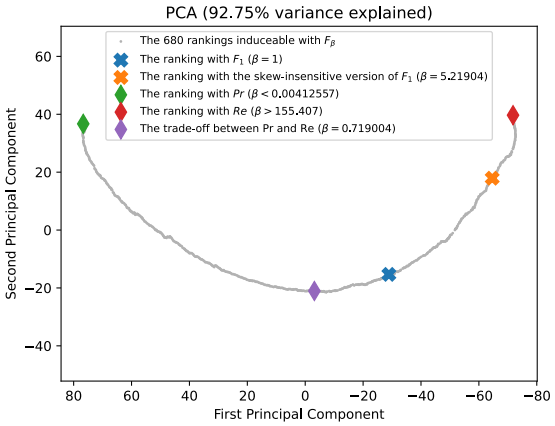
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



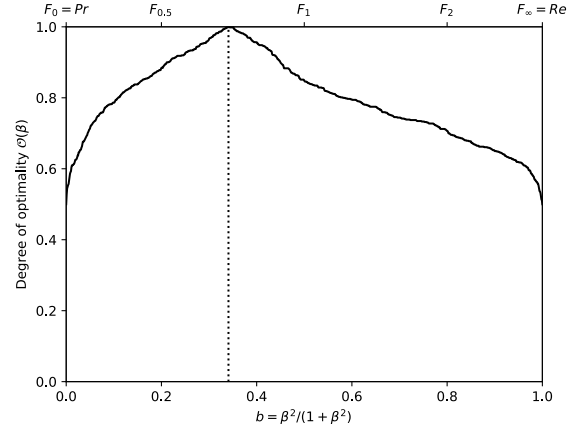
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

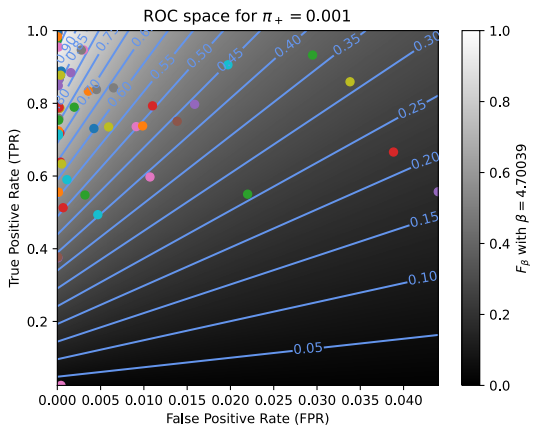


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

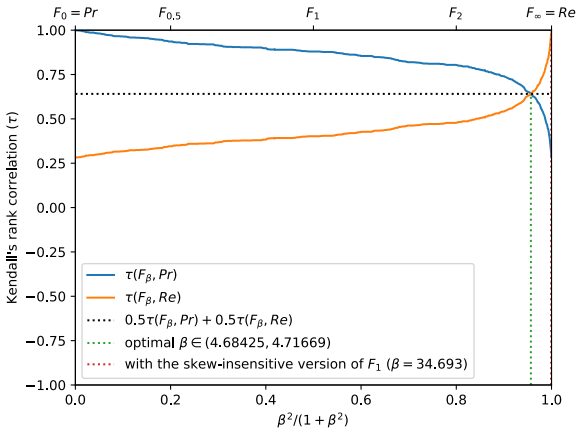


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

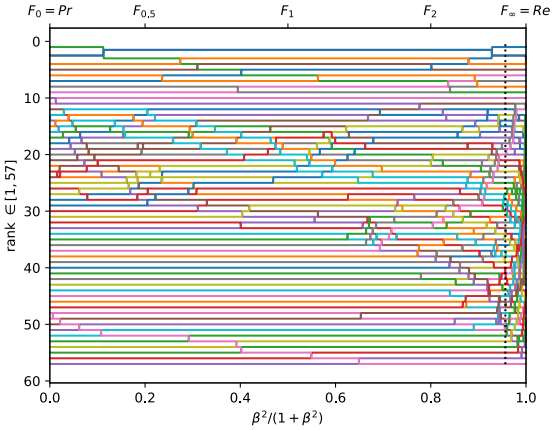
Figure A.3.15. Ranking of 57 BGS methods evaluated on the video "canoe" ($\pi_+ = 0.0354$).



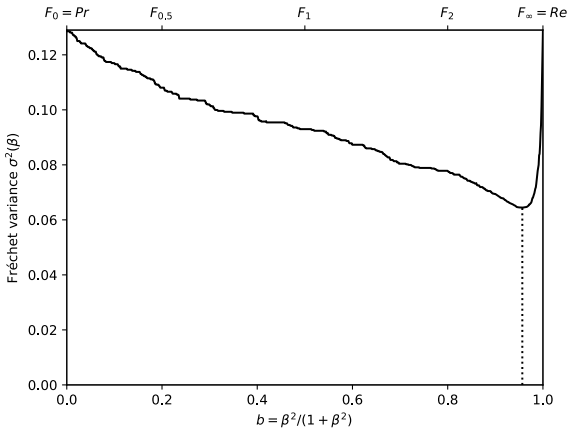
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



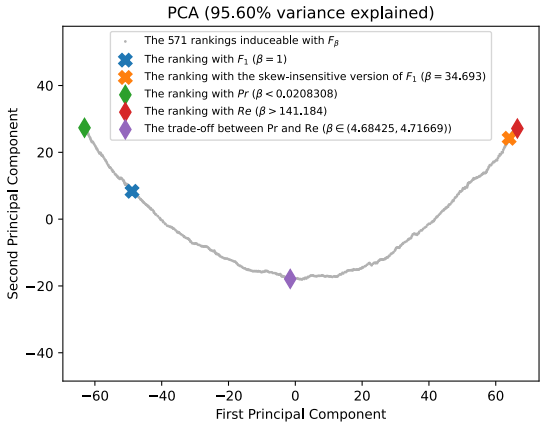
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



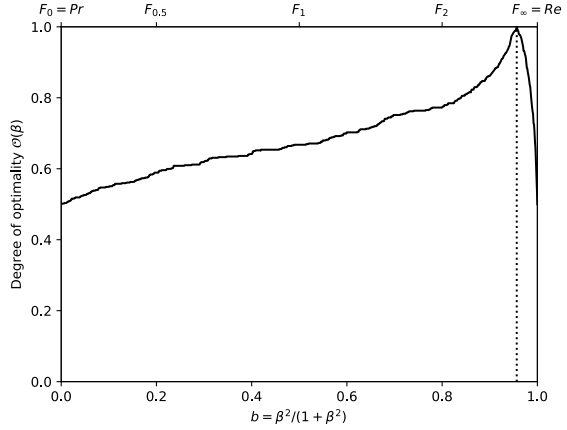
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

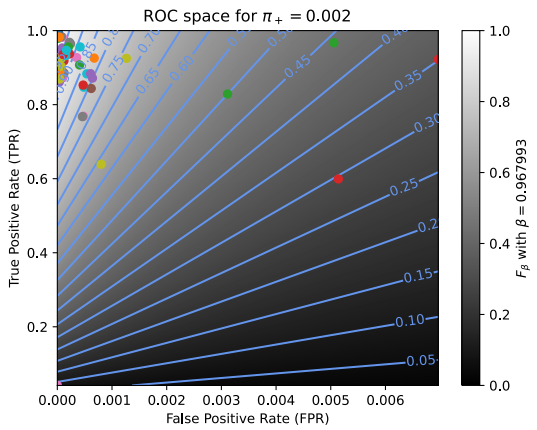


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

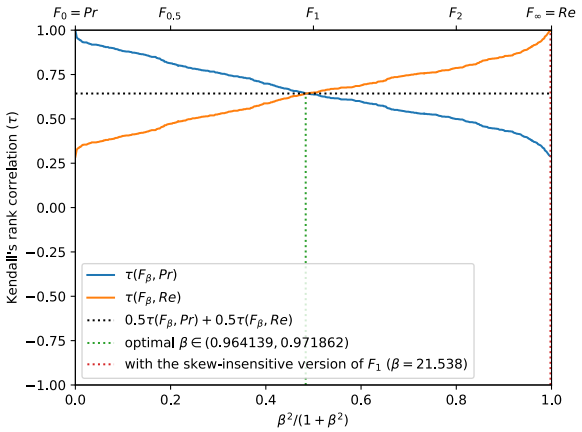


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

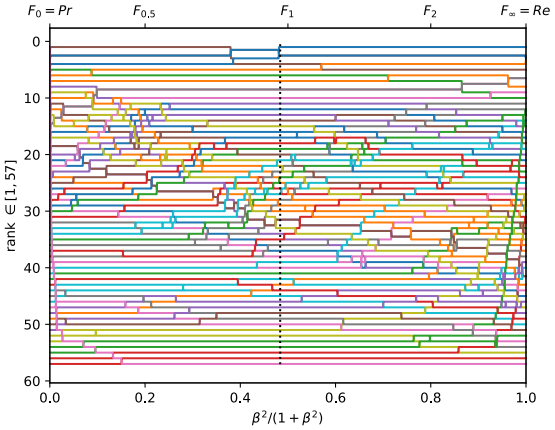
Figure A.3.16. Ranking of 57 BGS methods evaluated on the video "fountain01" ($\pi_+ = 0.0008$).



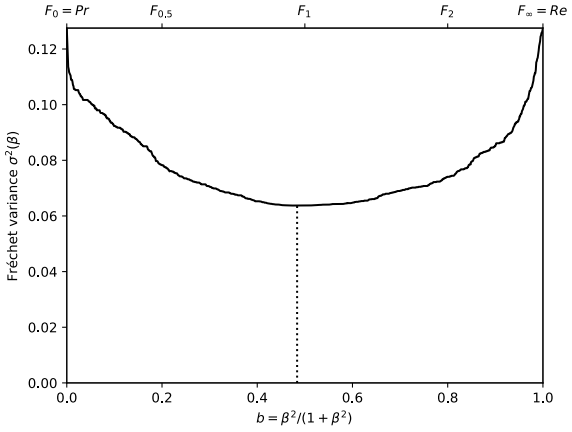
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



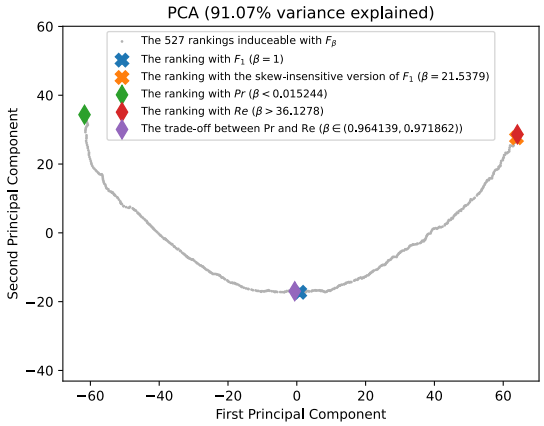
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



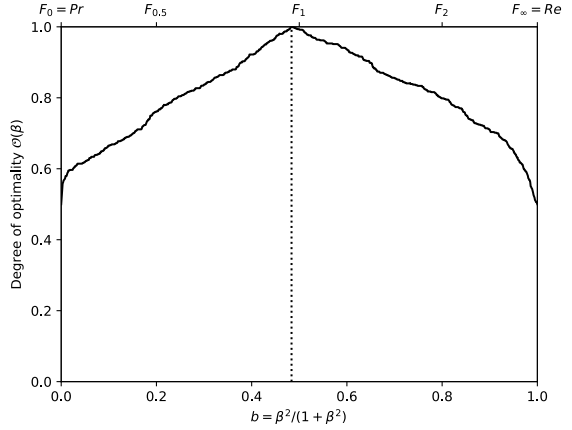
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

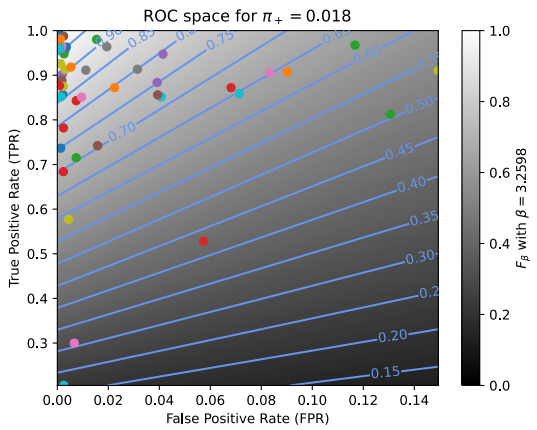


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

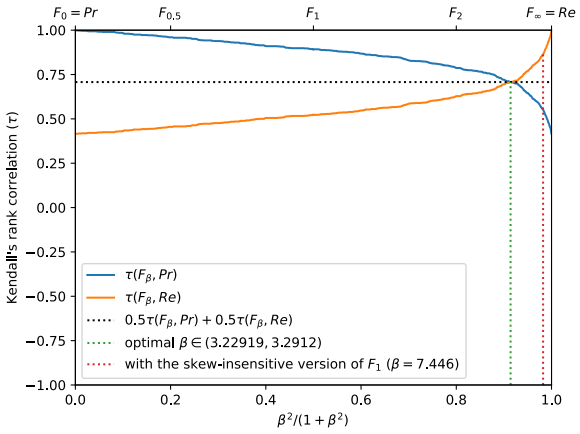


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

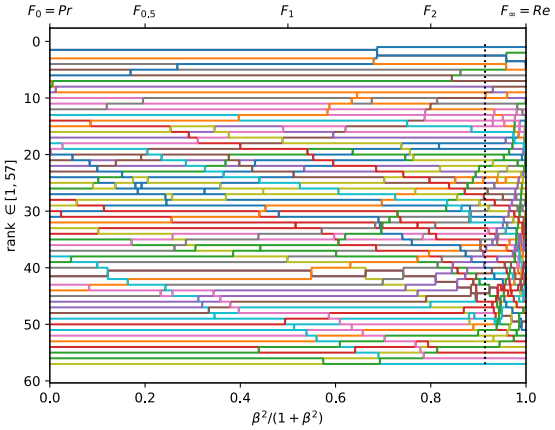
Figure A.3.17. Ranking of 57 BGS methods evaluated on the video "fountain02" ($\pi_+ = 0.0022$).



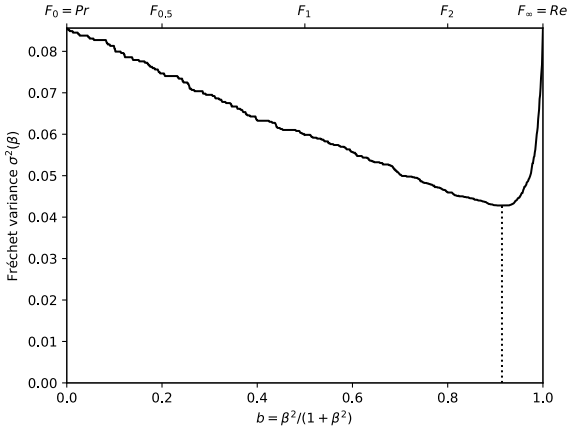
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



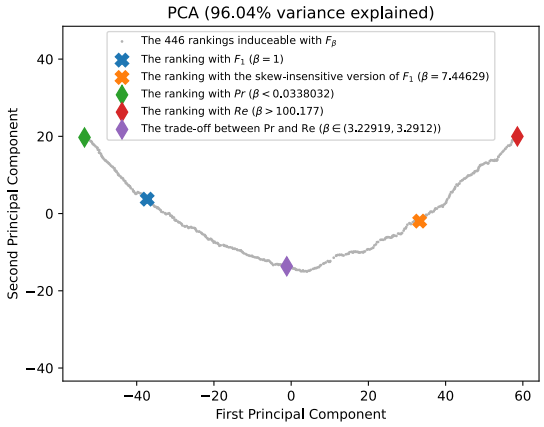
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



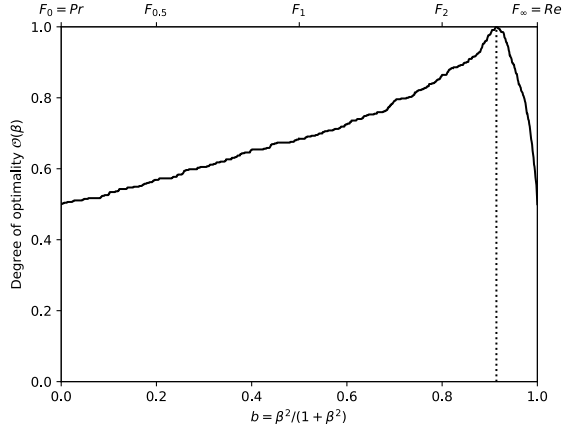
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

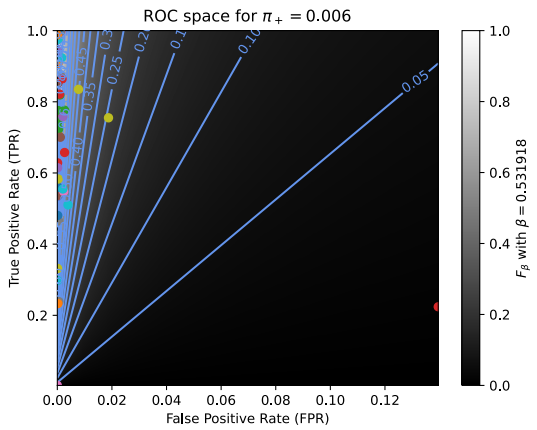


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

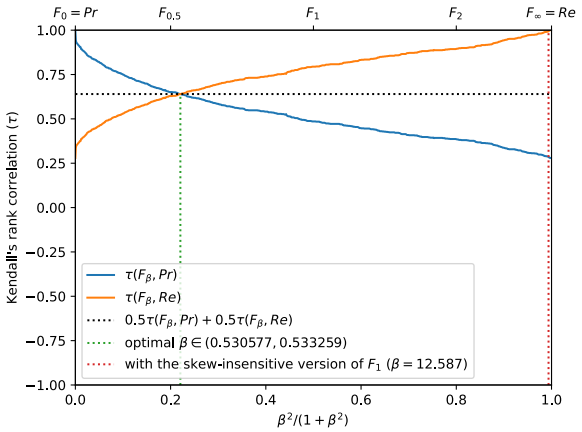


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

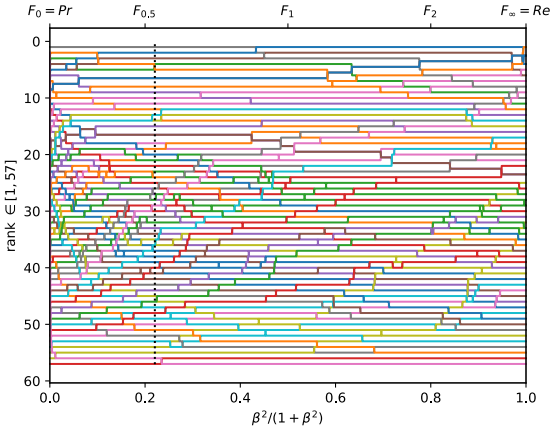
Figure A.3.18. Ranking of 57 BGS methods evaluated on the video "fall" ($\pi_+ = 0.0177$).



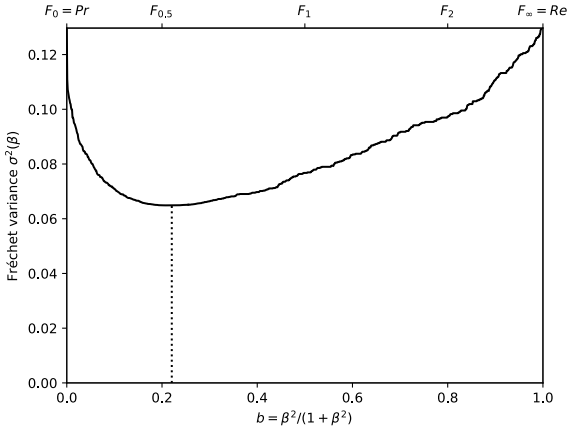
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



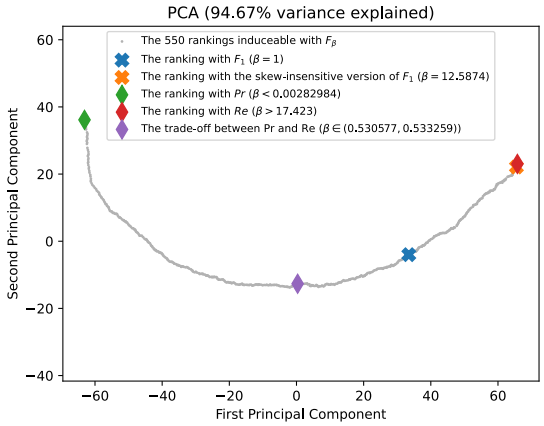
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



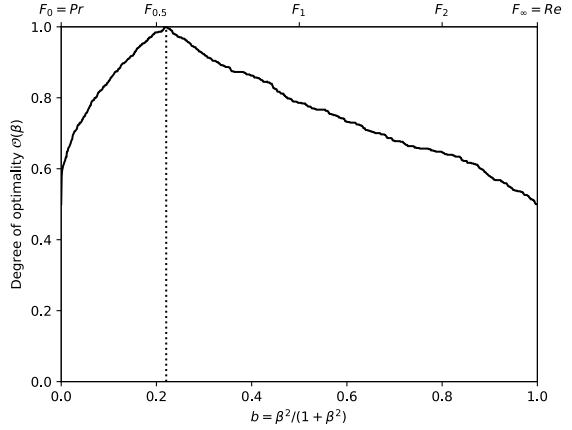
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

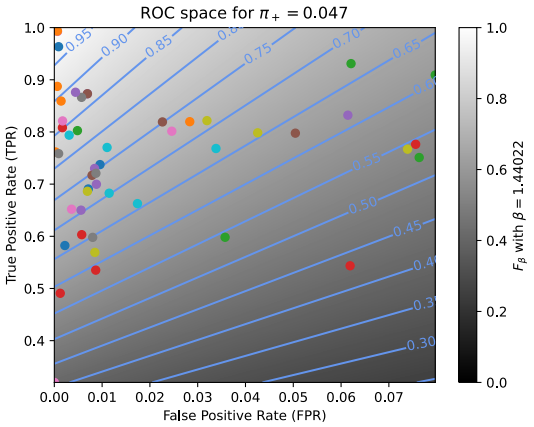


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

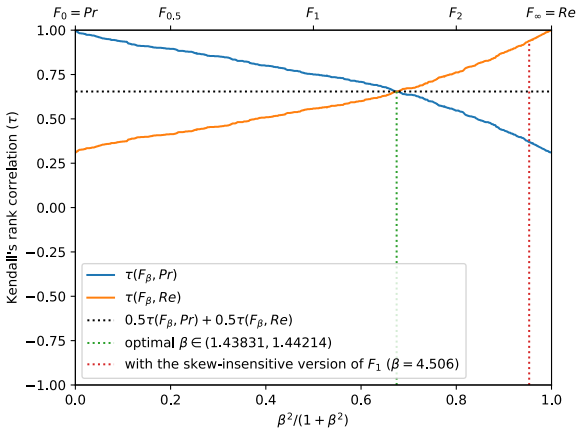


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

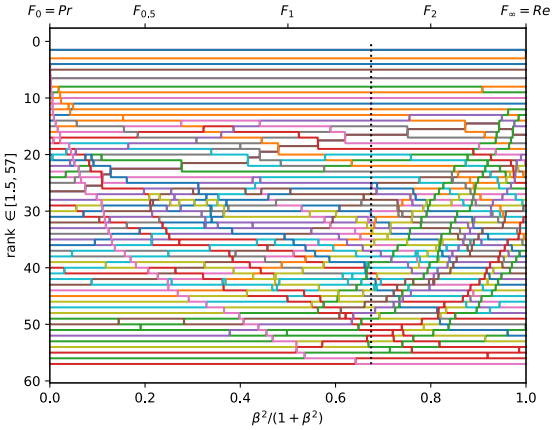
Figure A.3.19. Ranking of 57 BGS methods evaluated on the video "boats" ($\pi_+ = 0.0063$).



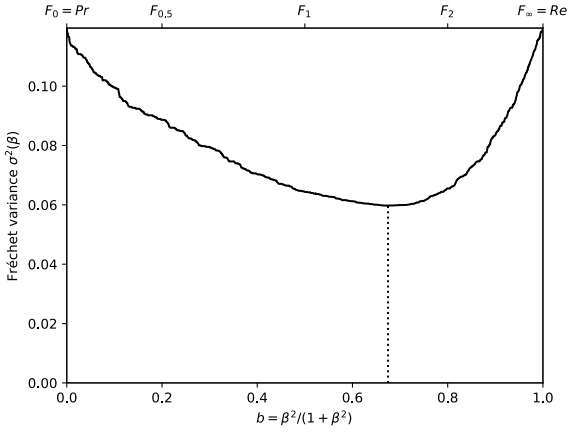
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



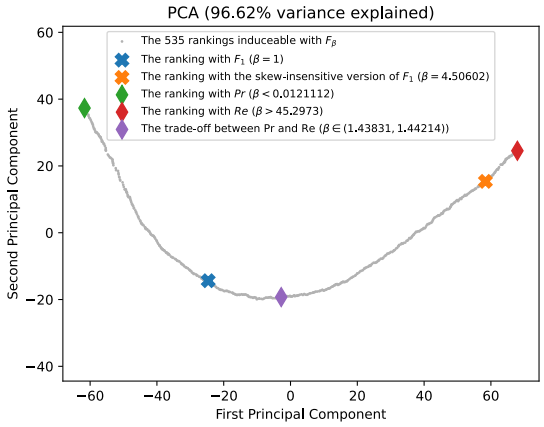
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



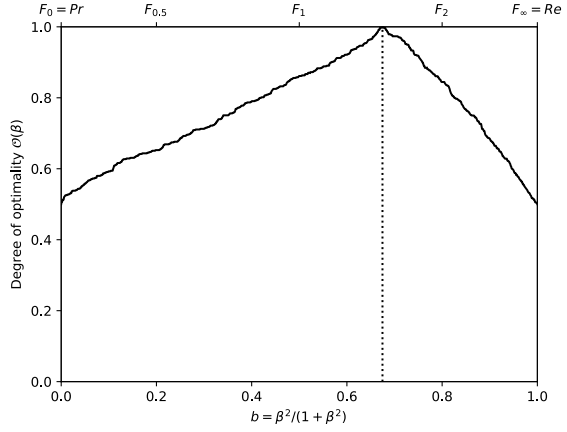
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

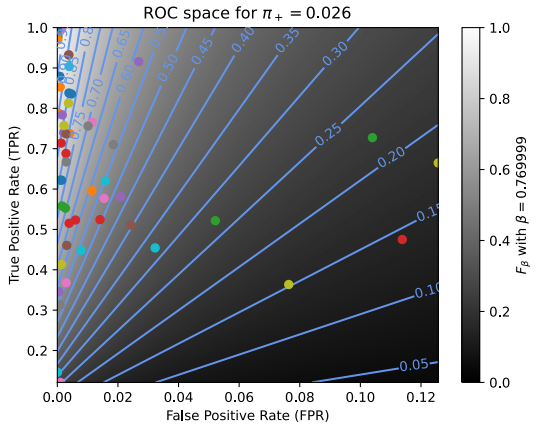


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

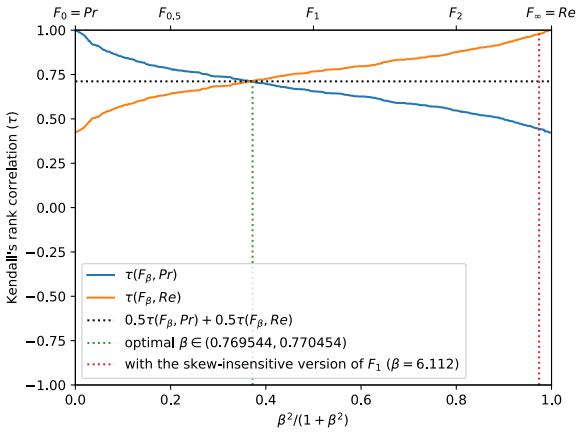


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

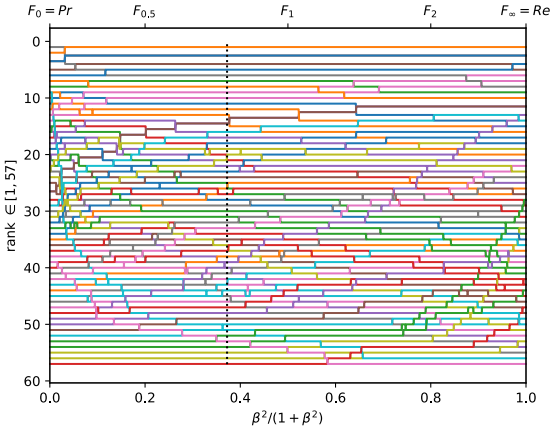
Figure A.3.20. Ranking of 57 BGS methods evaluated on the video "boulevard" ($\pi_+ = 0.0469$).



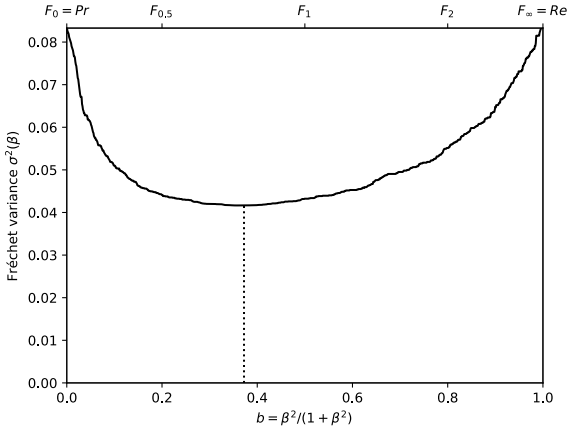
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



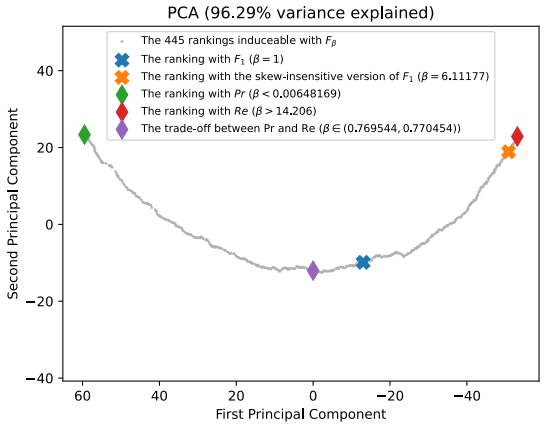
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



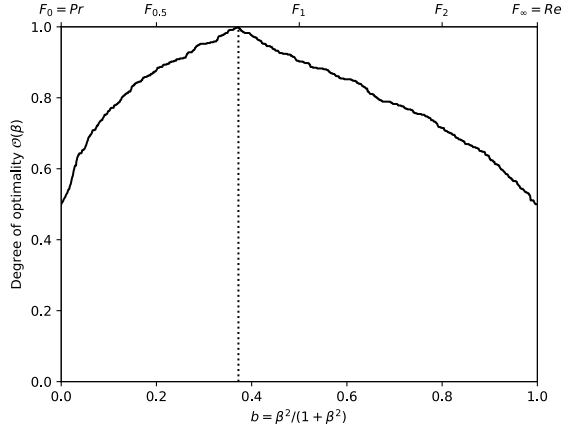
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

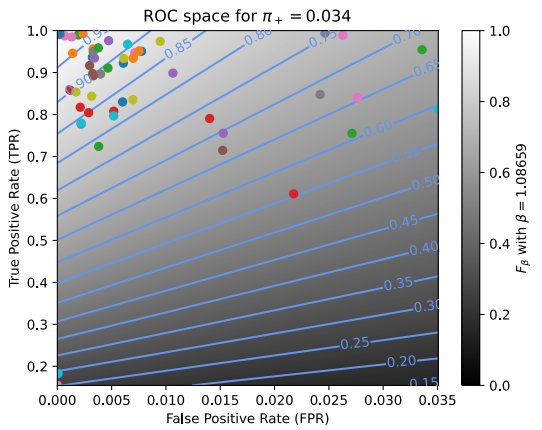


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

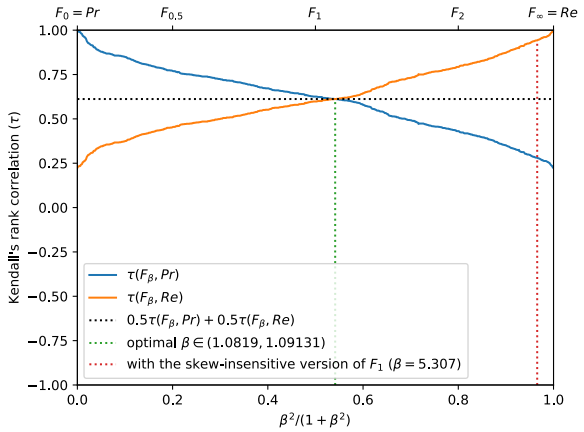


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

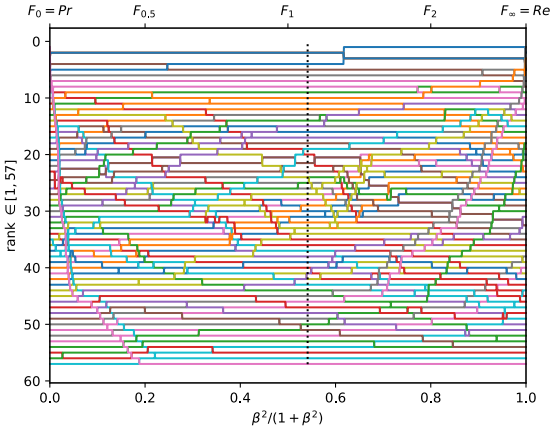
Figure A.3.21. Ranking of 57 BGS methods evaluated on the video "sidewalk" ($\pi_+ = 0.0261$).



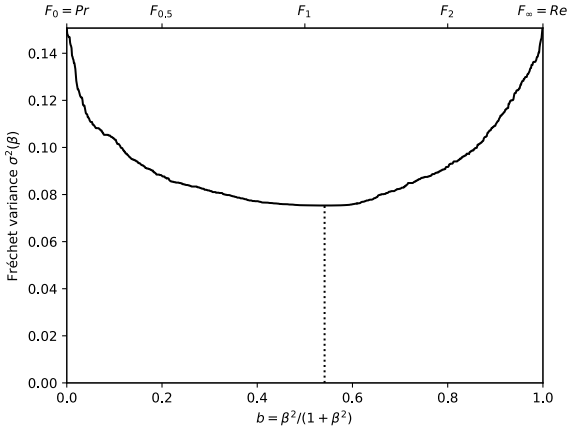
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



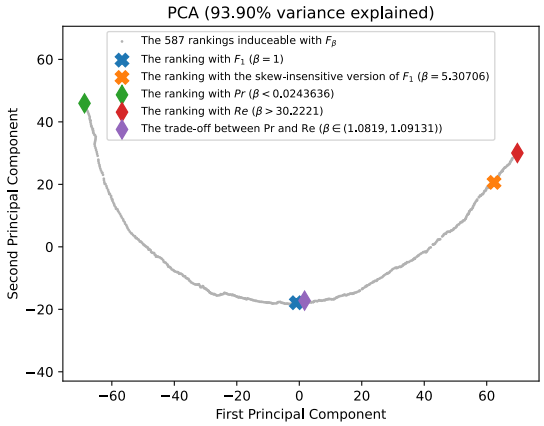
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



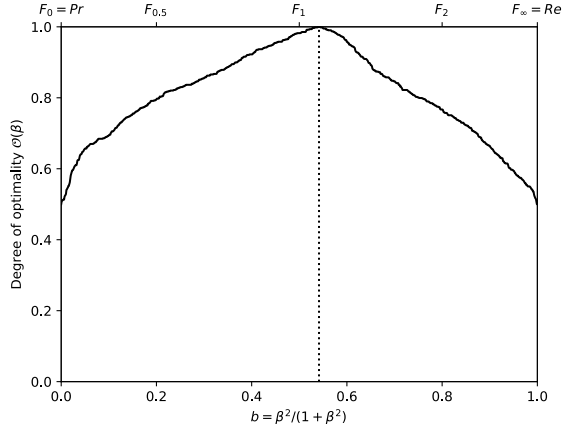
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

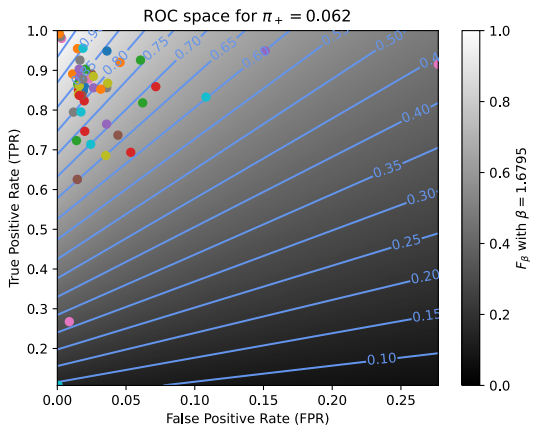


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

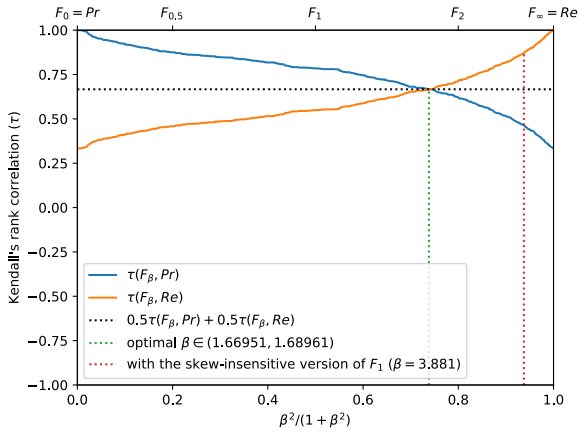


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

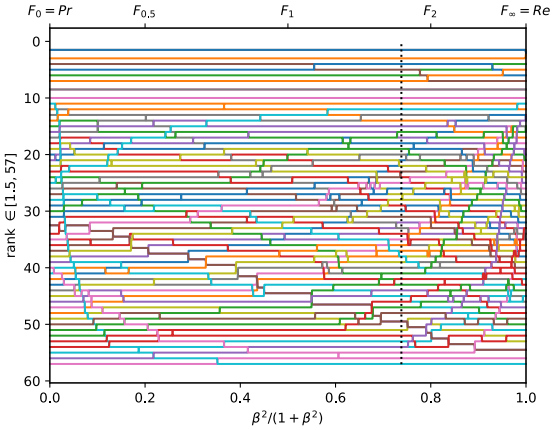
Figure A.3.22. Ranking of 57 BGS methods evaluated on the video "badminton" ($\pi_+ = 0.0343$).



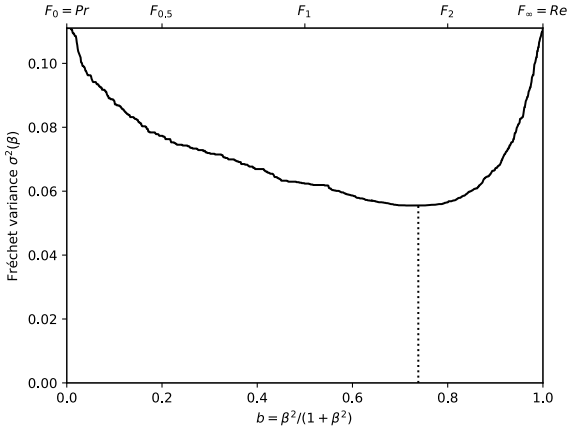
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



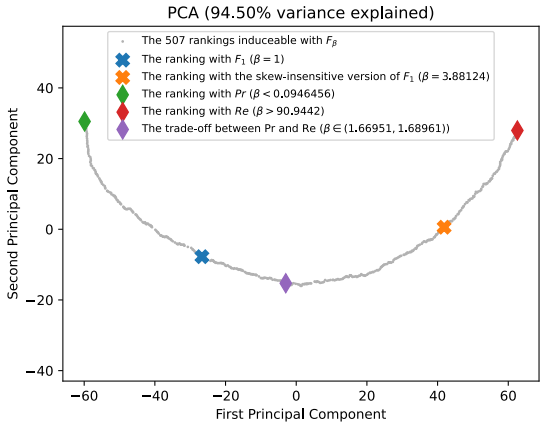
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



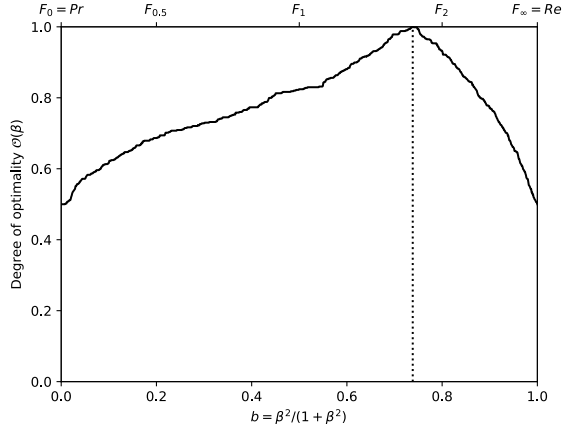
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

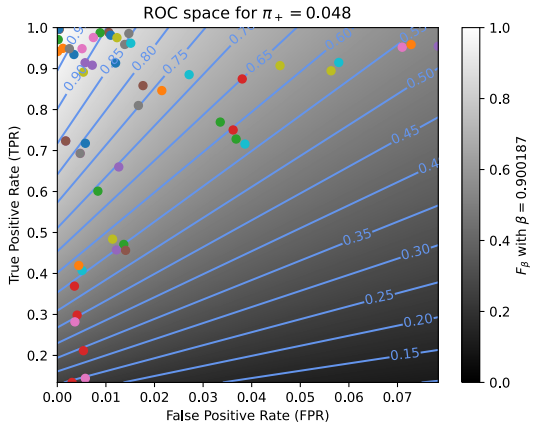


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

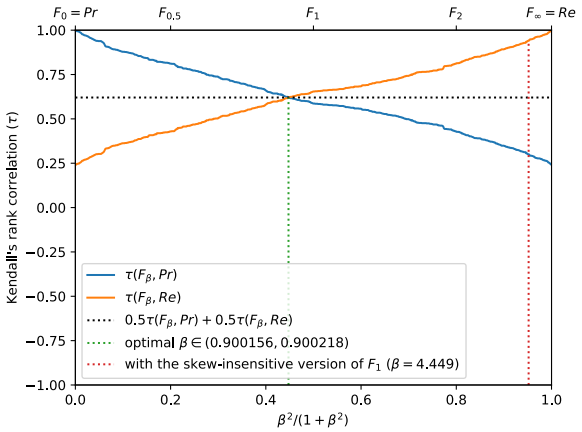


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

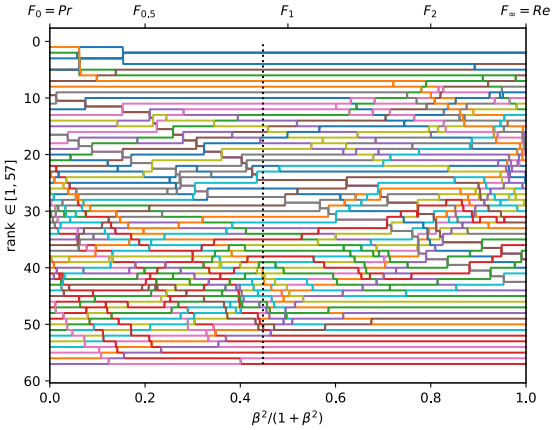
Figure A.3.23. Ranking of 57 BGS methods evaluated on the video "traffic" ($\pi_+ = 0.0623$).



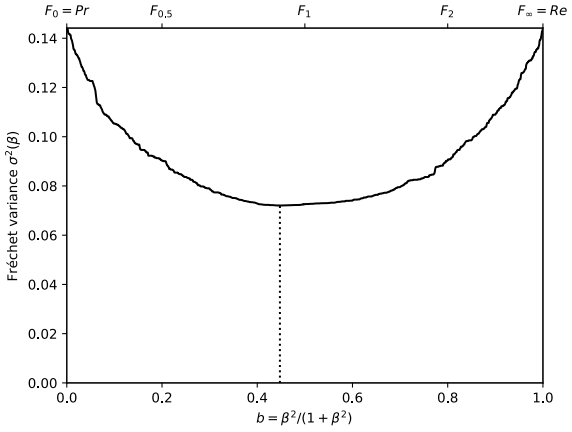
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



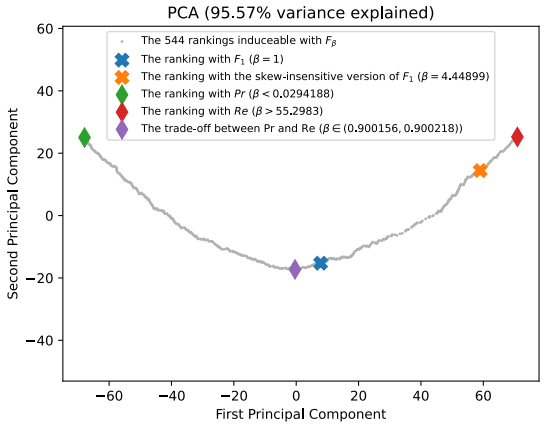
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



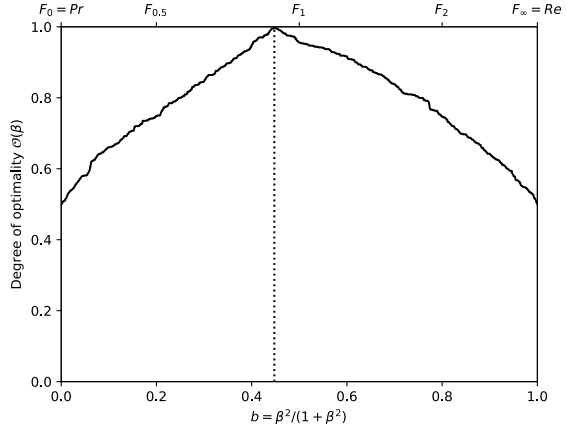
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

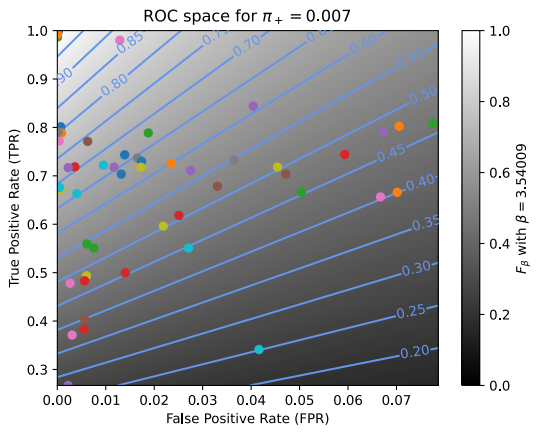


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

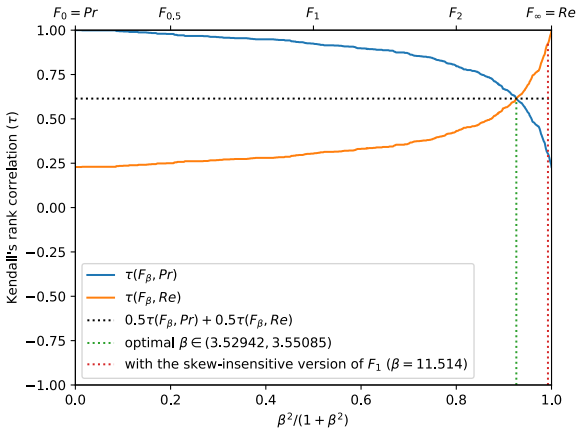


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

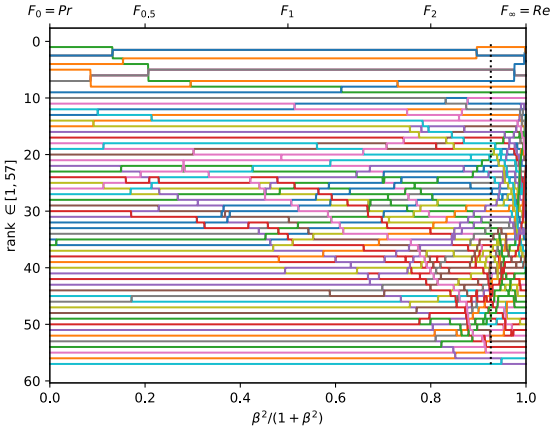
Figure A.3.24. Ranking of 57 BGS methods evaluated on the video "abandonedBox" ($\pi_+ = 0.0481$).



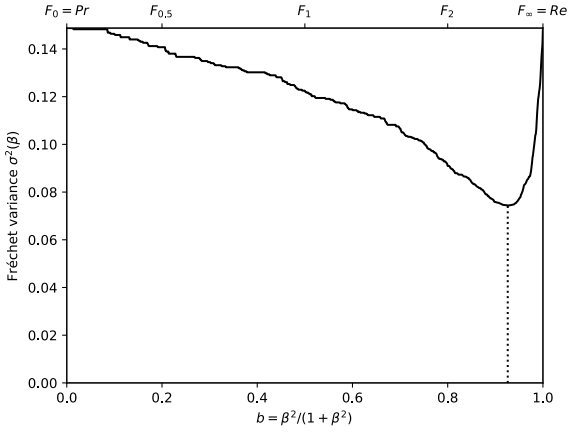
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



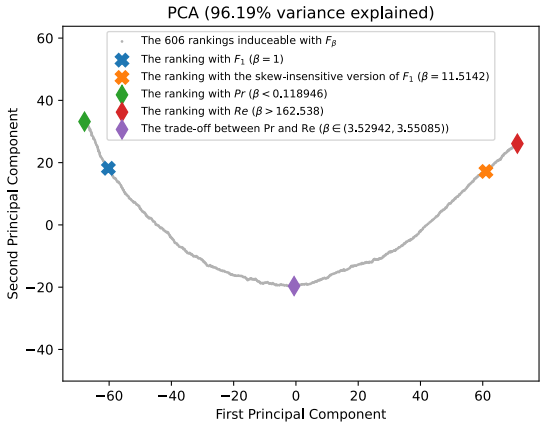
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



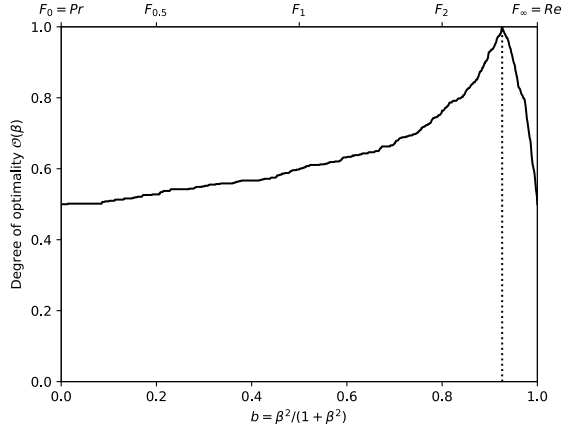
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

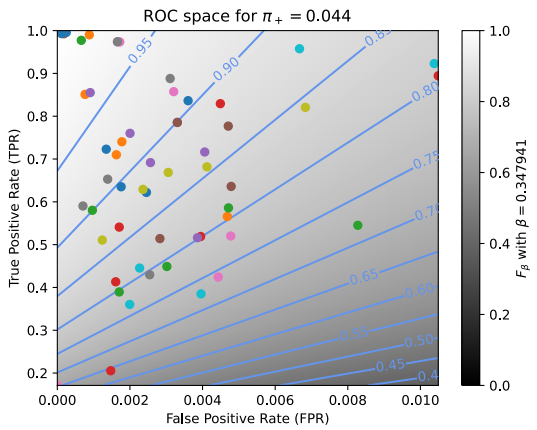


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

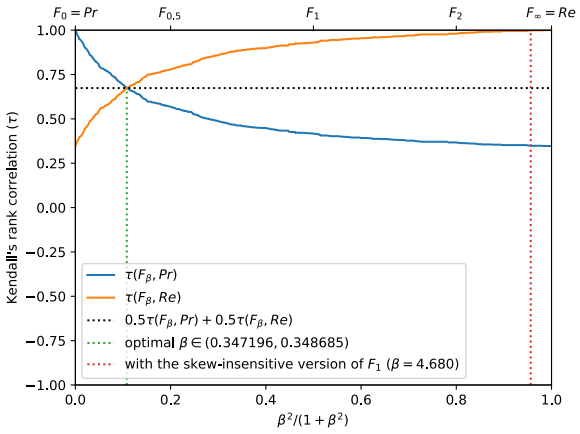


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

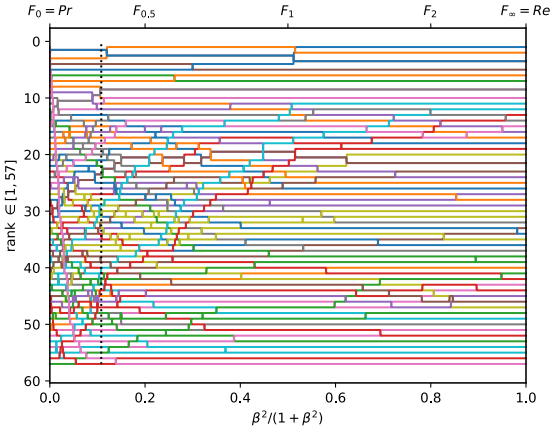
Figure A.3.25. Ranking of 57 BGS methods evaluated on the video "winterDriveway" ($\pi_+ = 0.0075$).



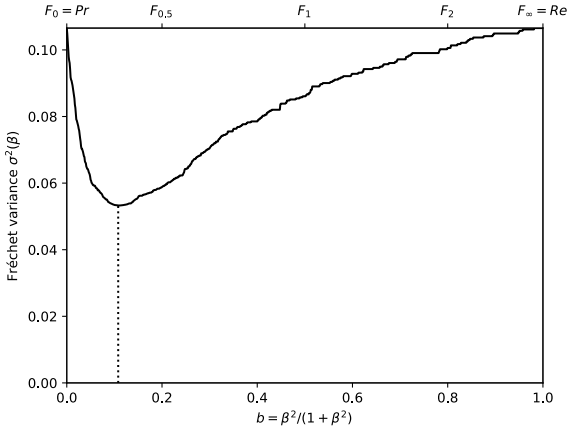
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



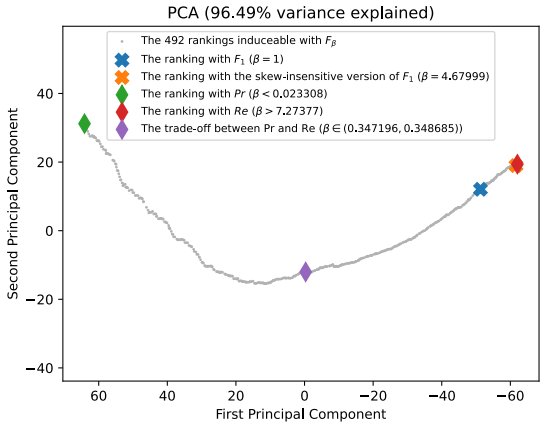
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



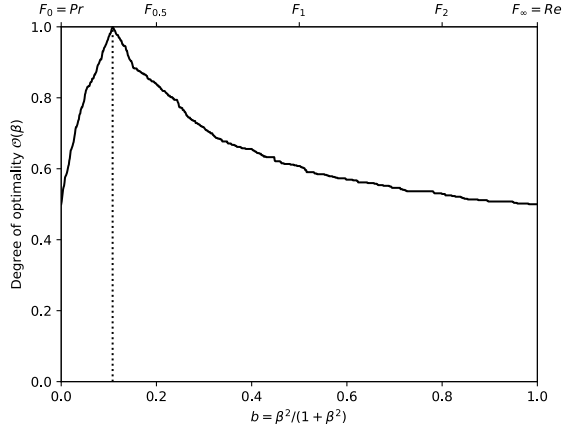
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

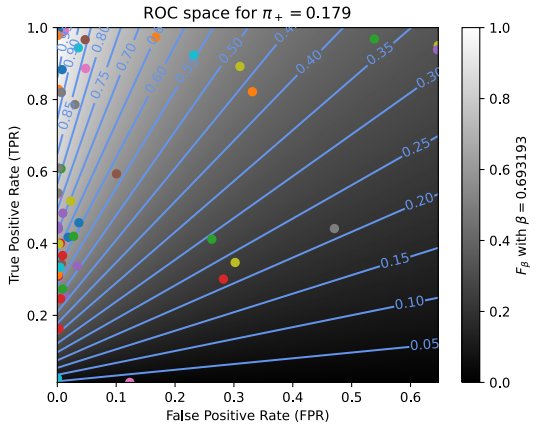


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

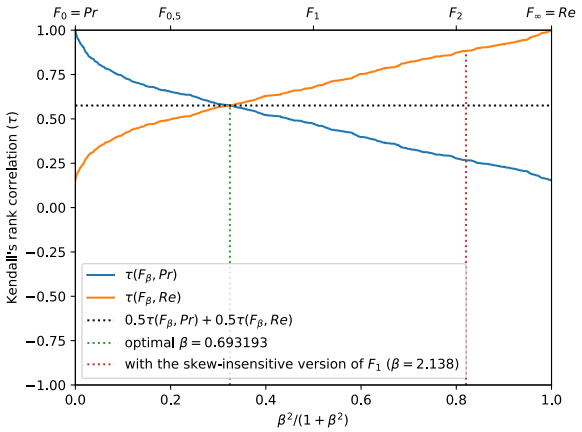


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

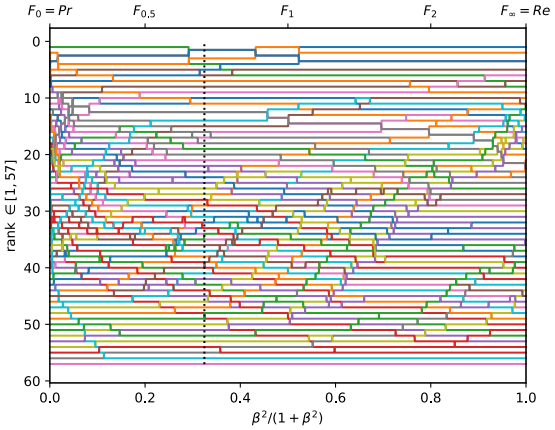
Figure A.3.26. Ranking of 57 BGS methods evaluated on the video "sofa" ($\pi_+ = 0.0437$).



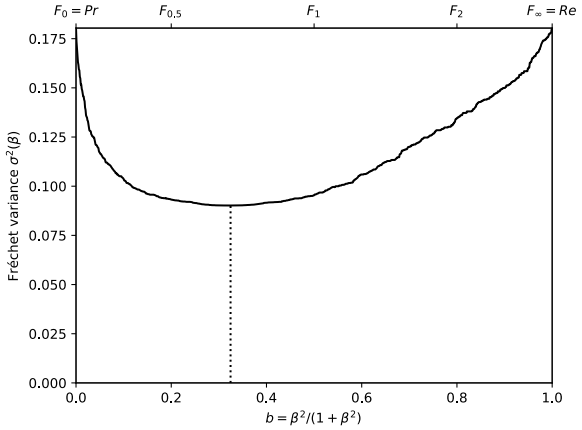
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



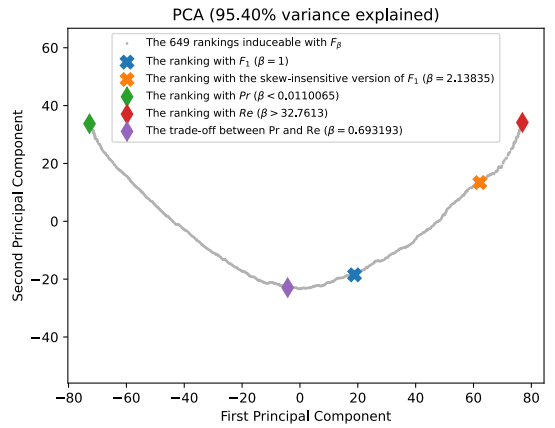
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



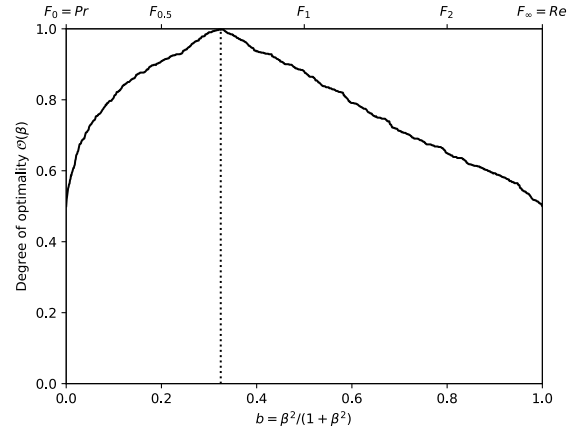
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

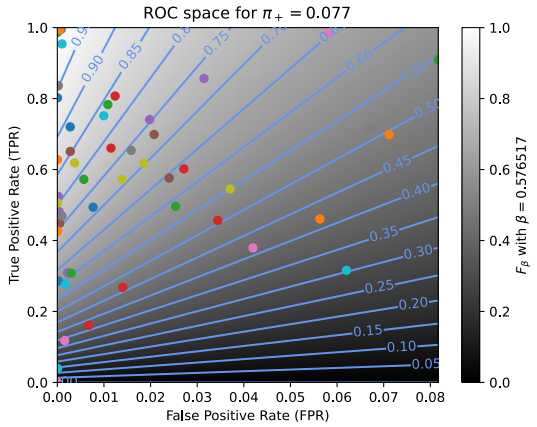


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

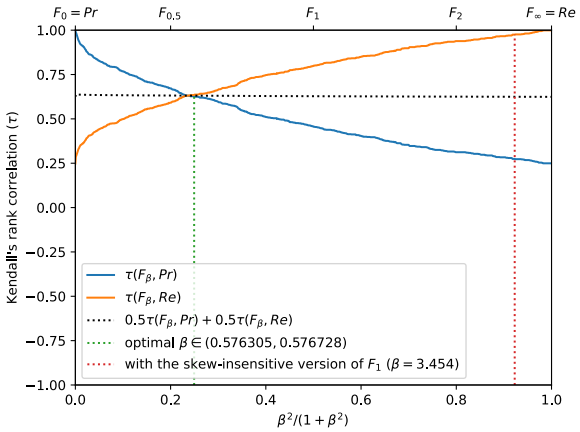


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

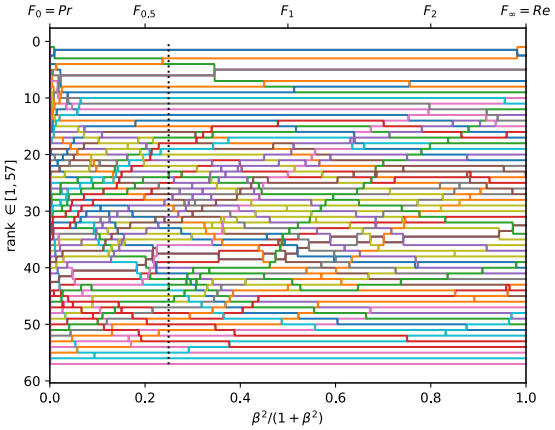
Figure A.3.27. Ranking of 57 BGS methods evaluated on the video "tramstop" ($\pi_+ = 0.1795$).



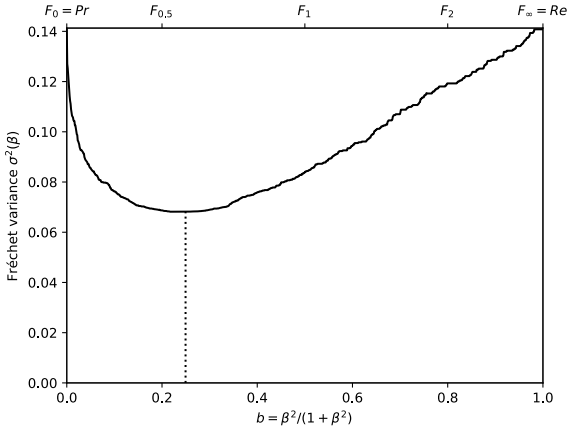
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



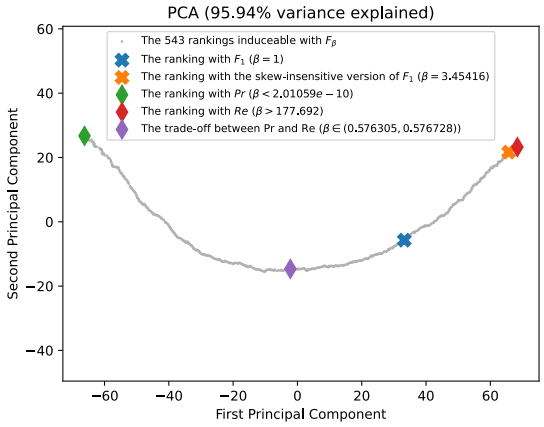
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



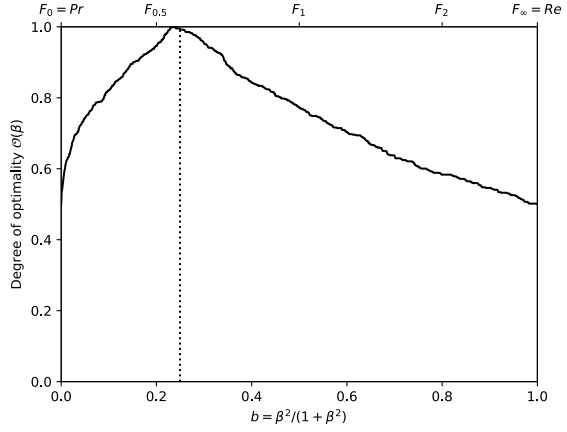
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

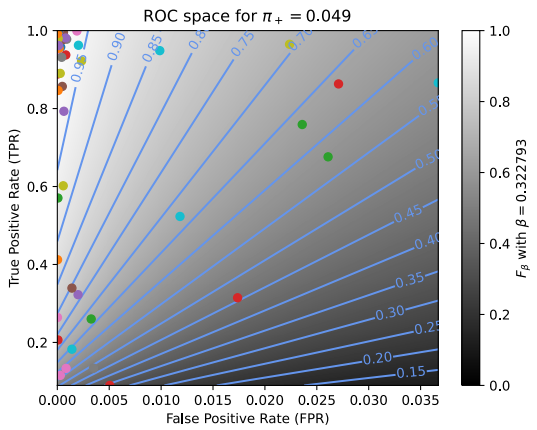


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

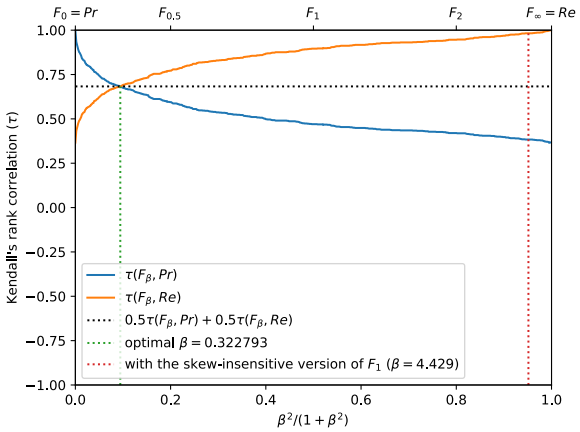


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

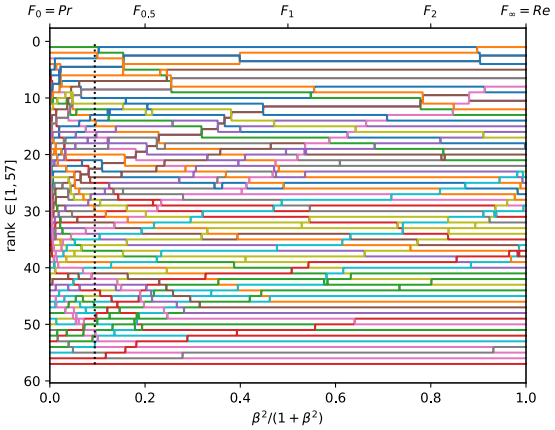
Figure A.3.28. Ranking of 57 BGS methods evaluated on the video "parking" ($\pi_+ = 0.0773$).



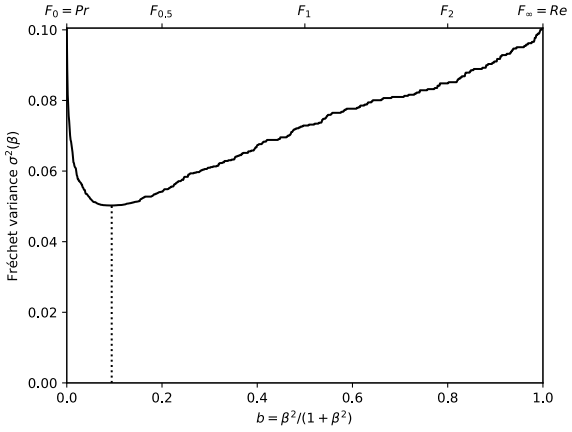
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



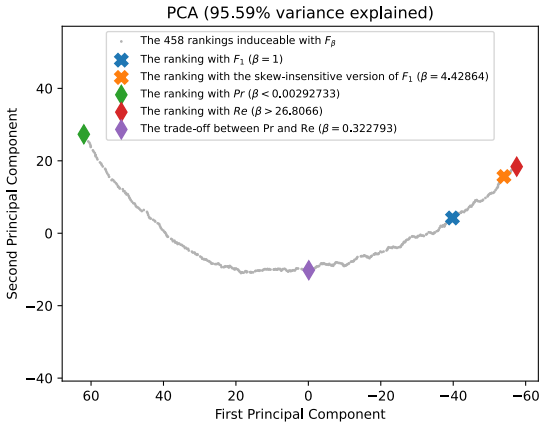
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



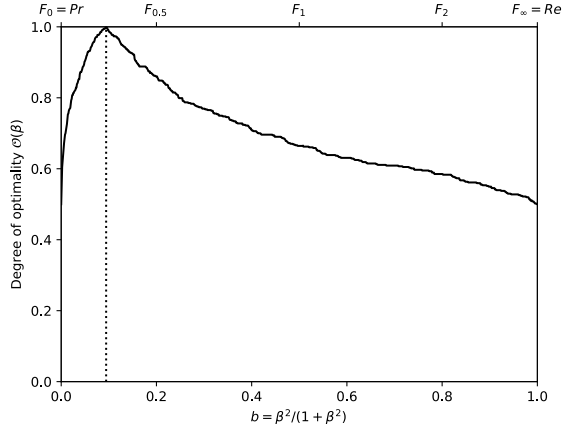
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

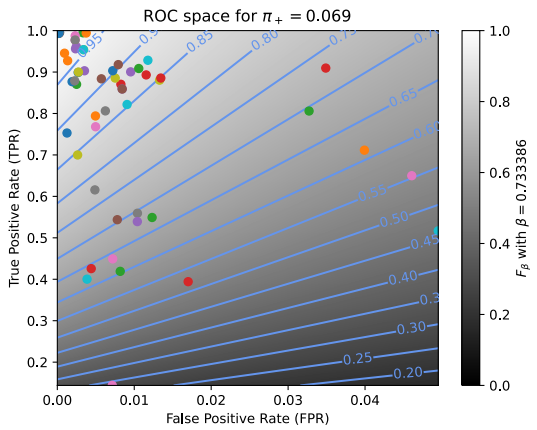


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

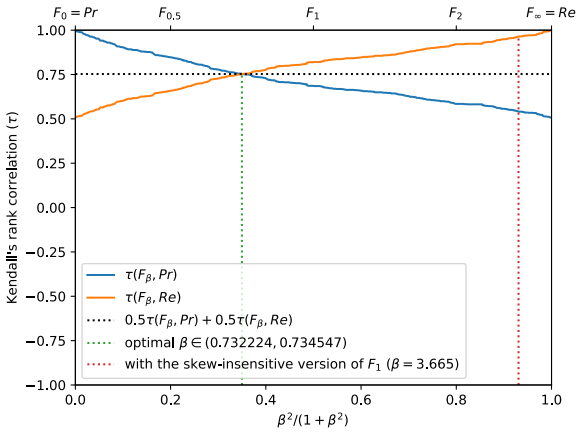


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

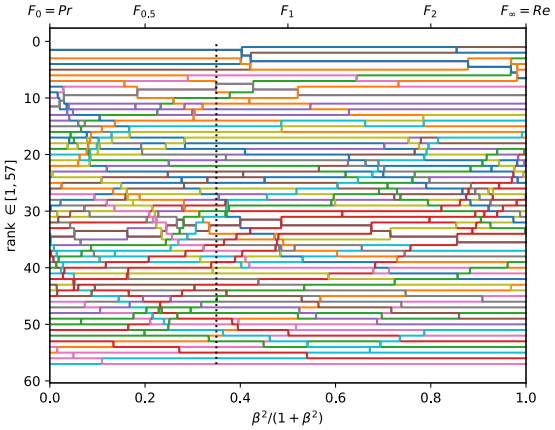
Figure A.3.29. Ranking of 57 BGS methods evaluated on the video "streetLight" ($\pi_+ = 0.0485$).



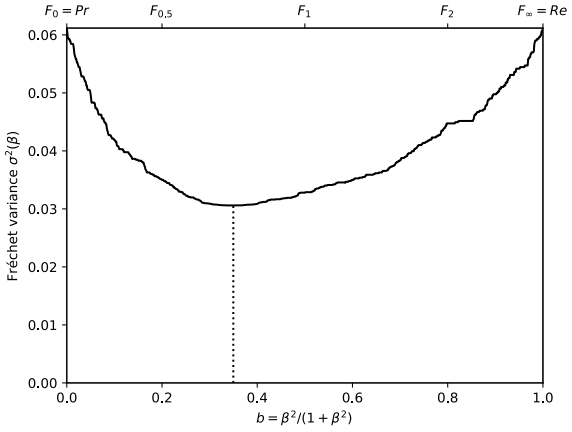
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



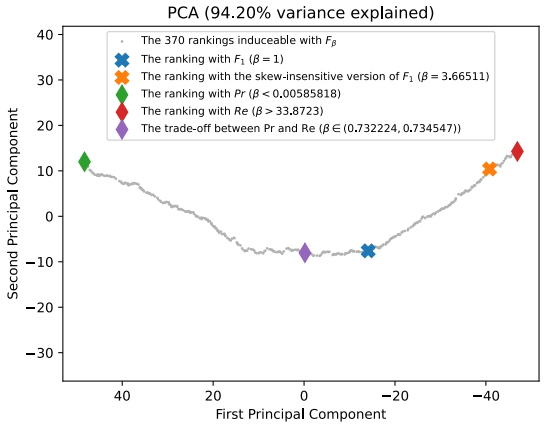
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



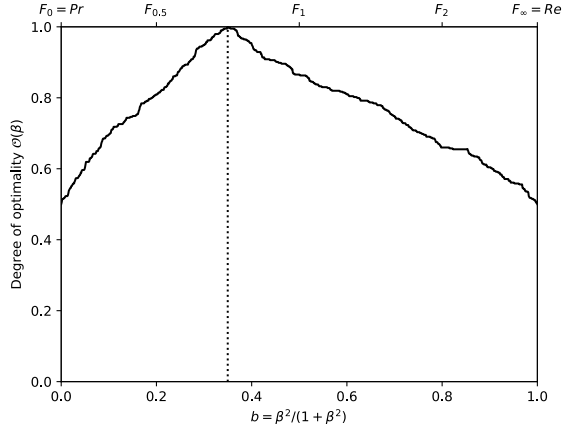
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

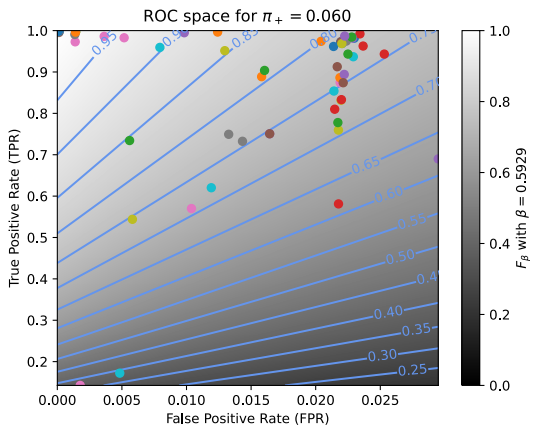


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

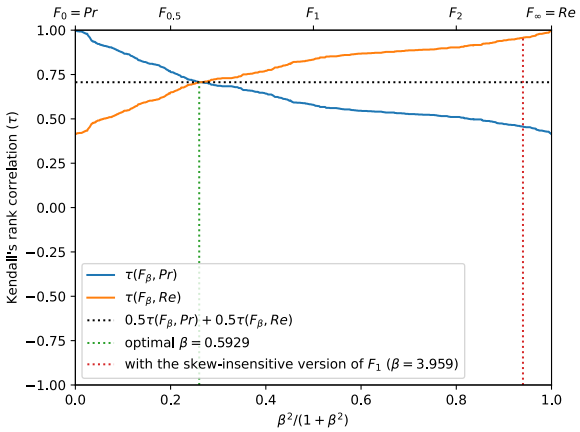


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

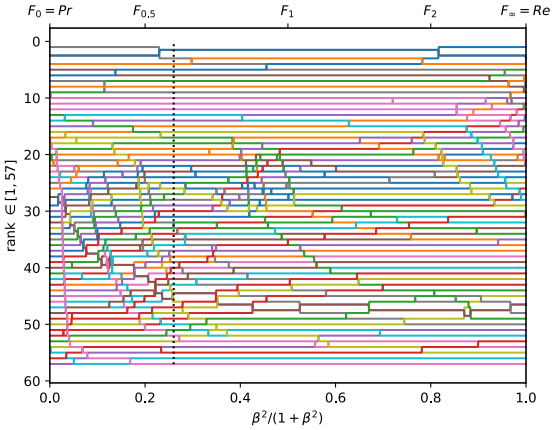
Figure A.3.30. Ranking of 57 BGS methods evaluated on the video "copyMachine" ($\pi_+ = 0.0693$).



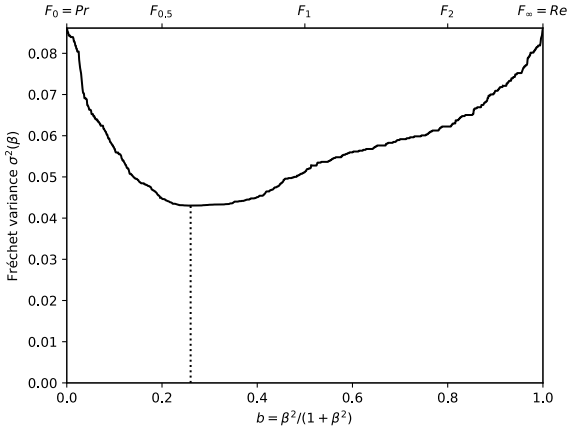
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



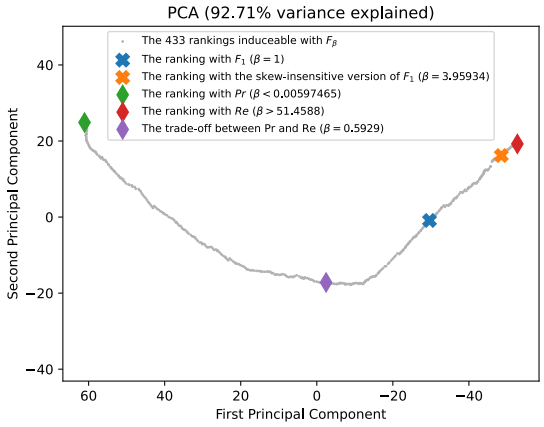
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



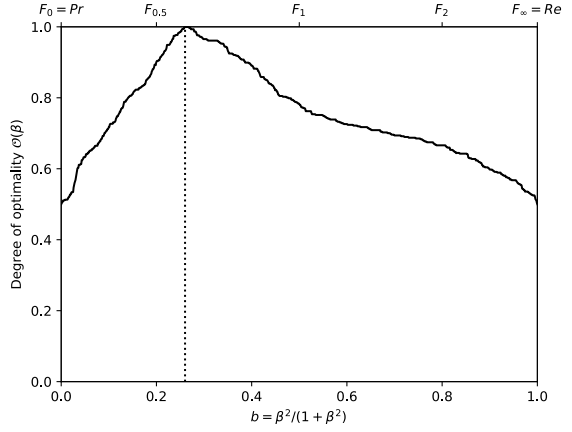
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

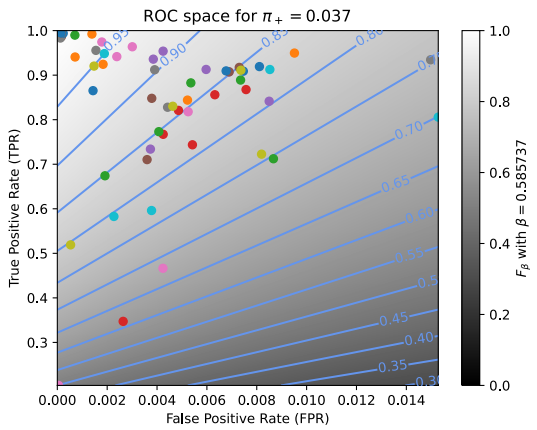


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

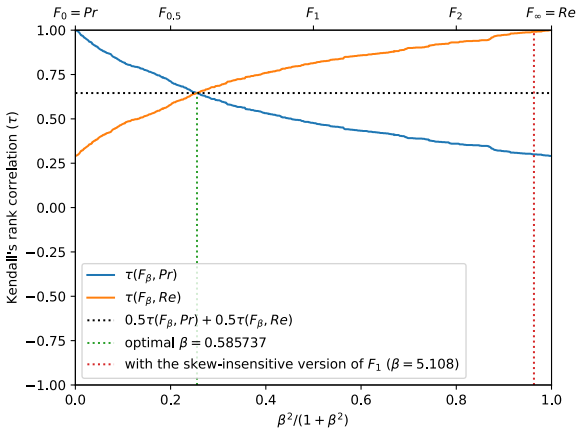


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

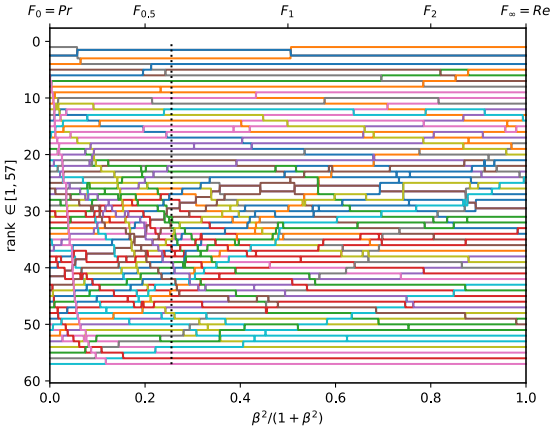
Figure A.3.31. Ranking of 57 BGS methods evaluated on the video "bungalows" ($\pi_+ = 0.0600$).



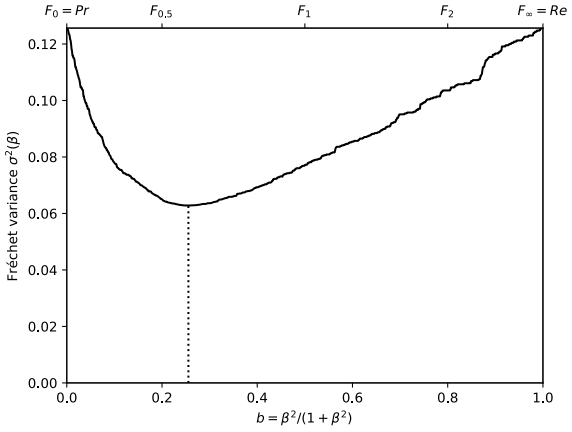
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



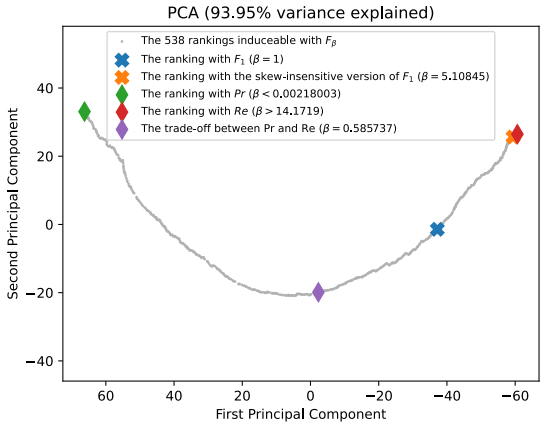
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



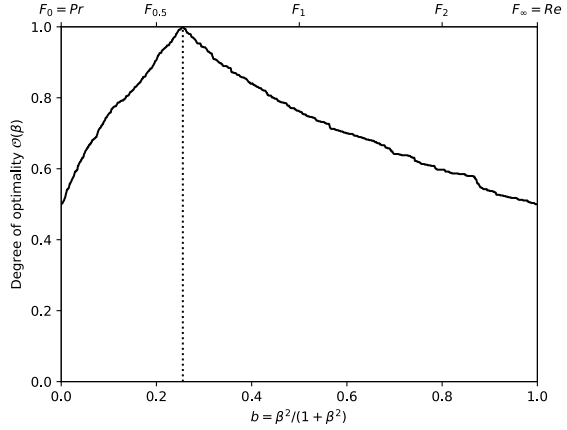
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

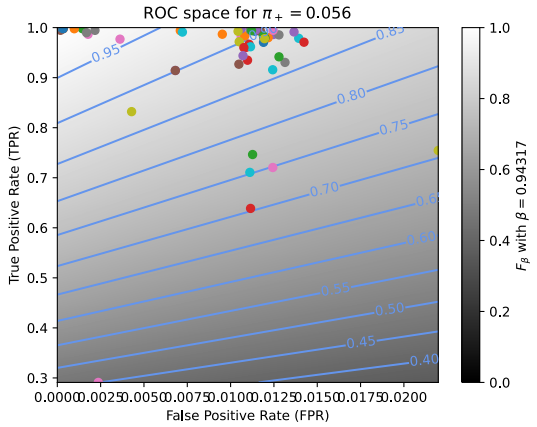


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

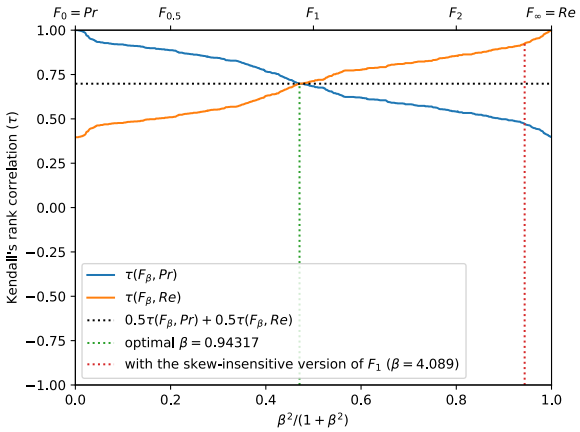


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

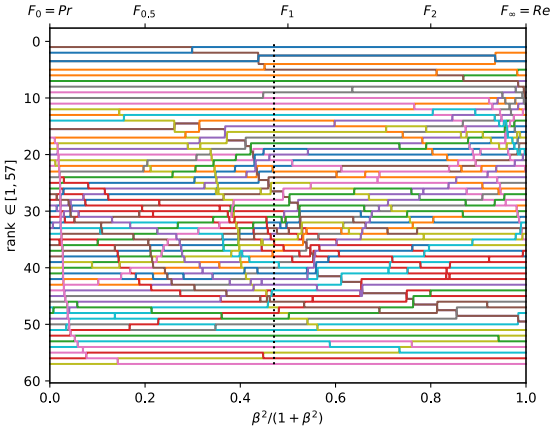
Figure A.3.32. Ranking of 57 BGS methods evaluated on the video "busStation" ($\pi_+ = 0.0369$).



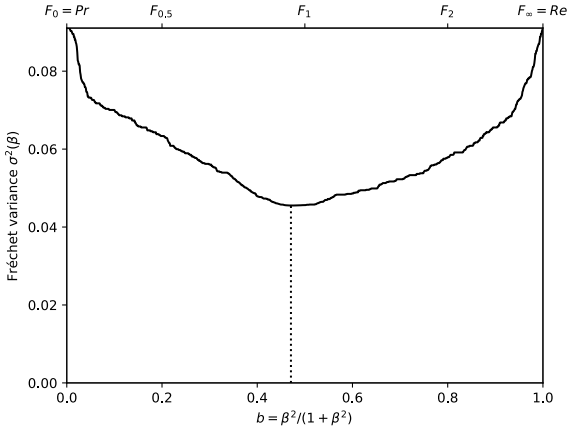
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



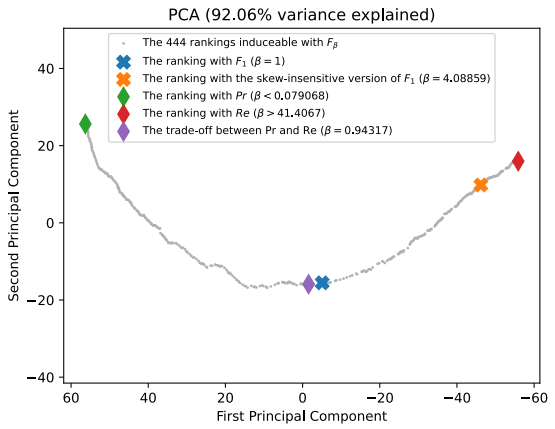
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



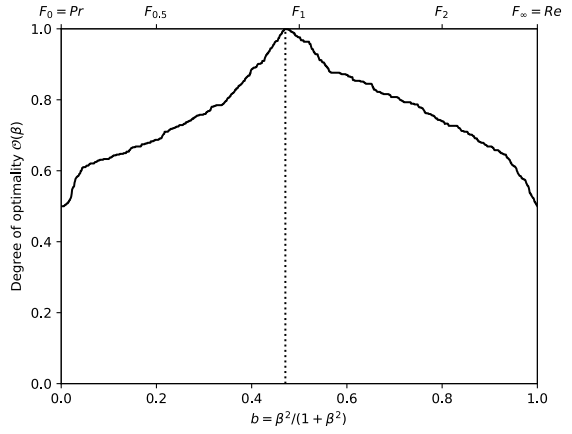
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

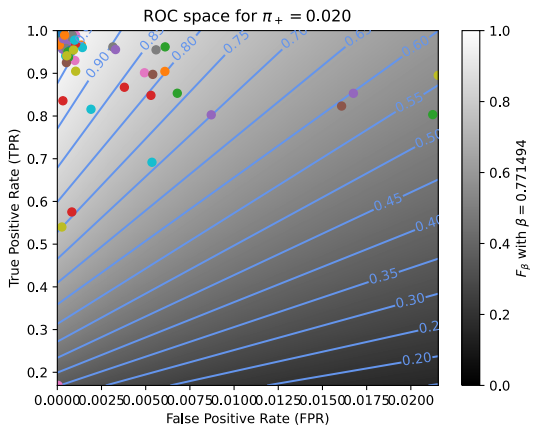


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

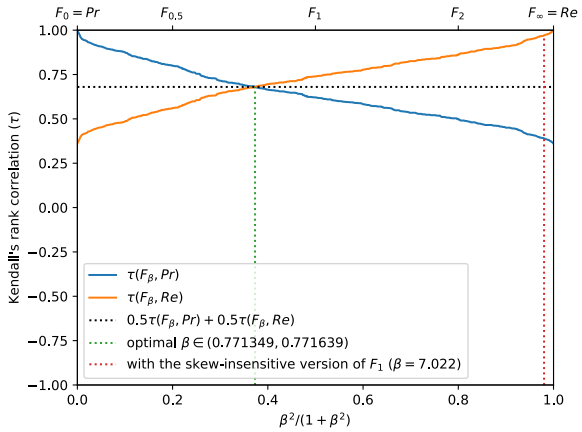


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

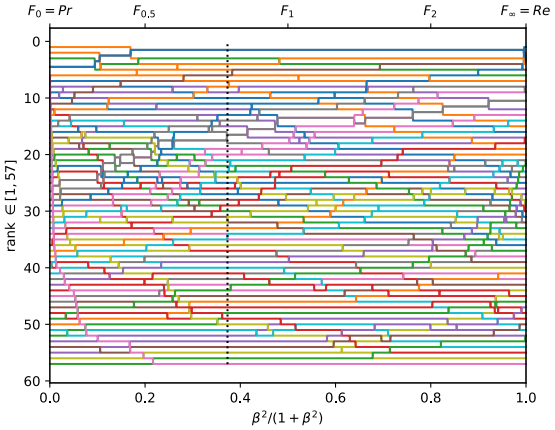
Figure A.3.33. Ranking of 57 BGS methods evaluated on the video "peopleInShade" ($\pi_+ = 0.0564$).



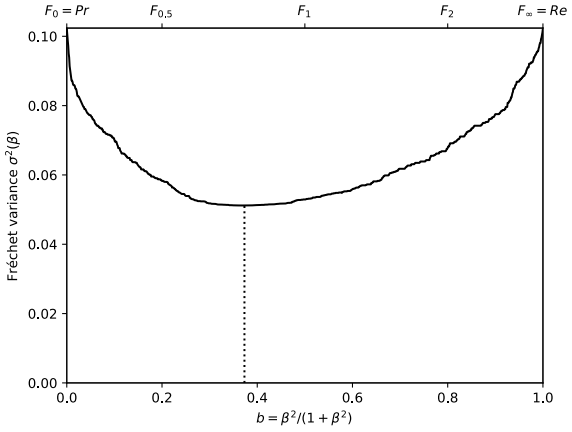
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



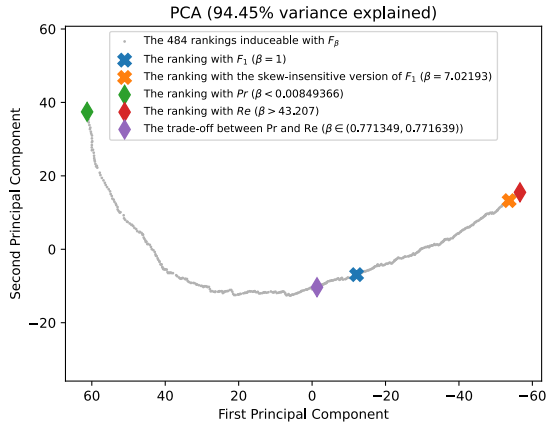
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



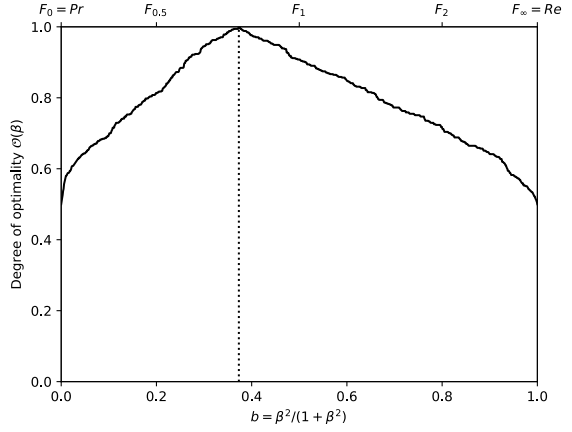
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

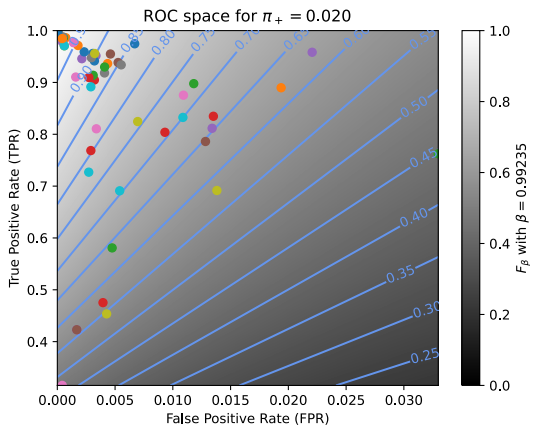


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

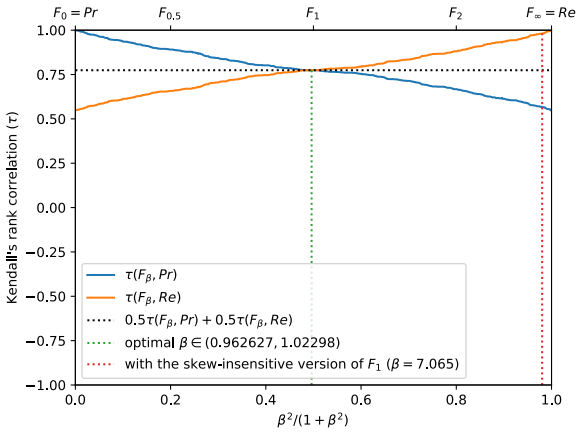


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

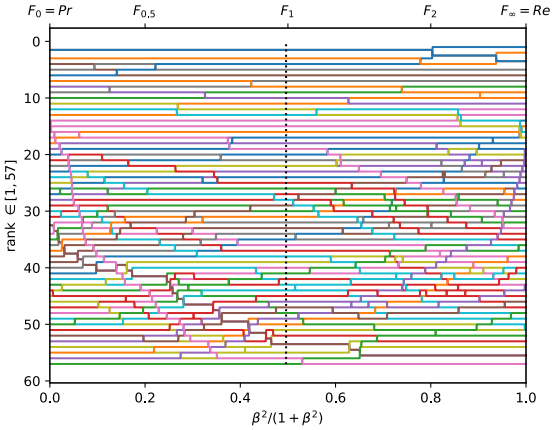
Figure A.3.34. Ranking of 57 BGS methods evaluated on the video "backdoor" ($\pi_+ = 0.0199$).



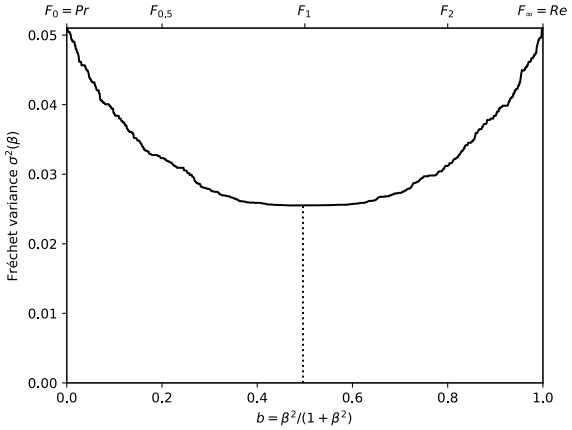
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



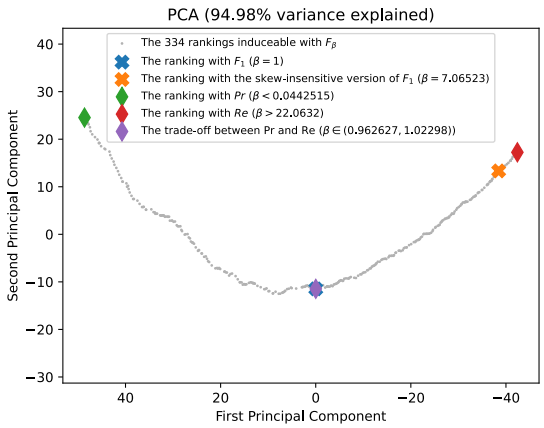
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



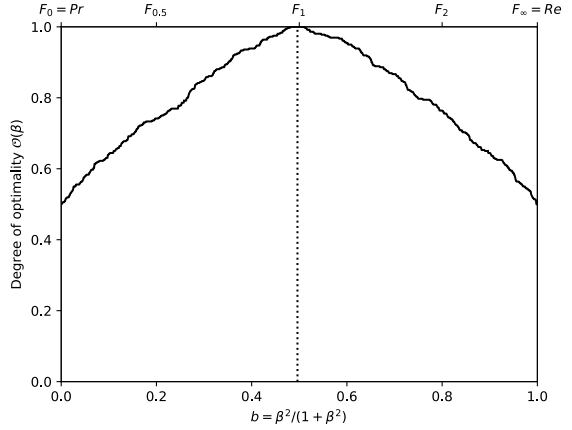
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

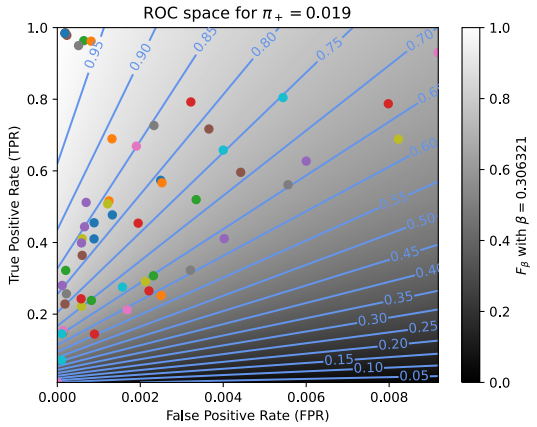


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

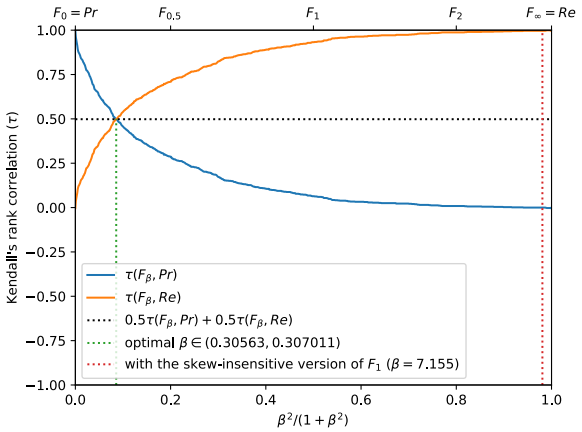


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

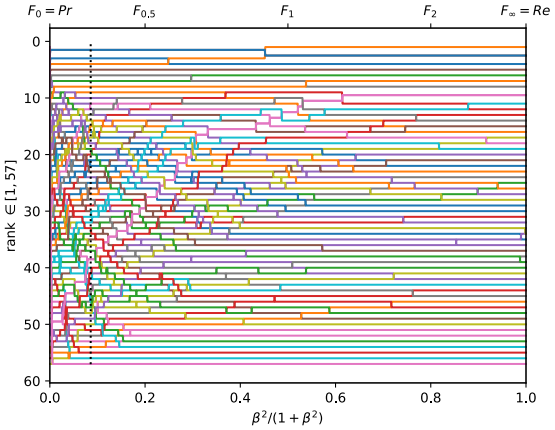
Figure A.3.35. Ranking of 57 BGS methods evaluated on the video "cubicle" ($\pi_+ = 0.0196$).



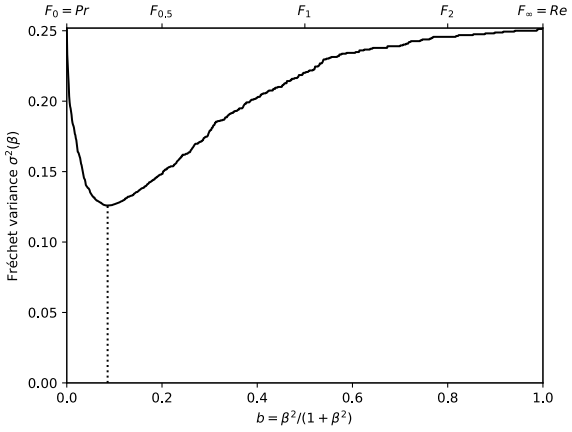
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



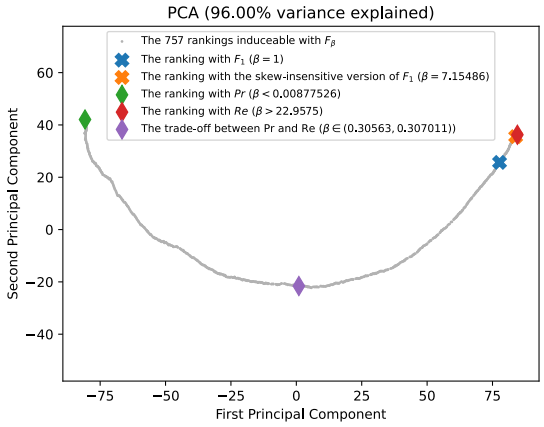
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



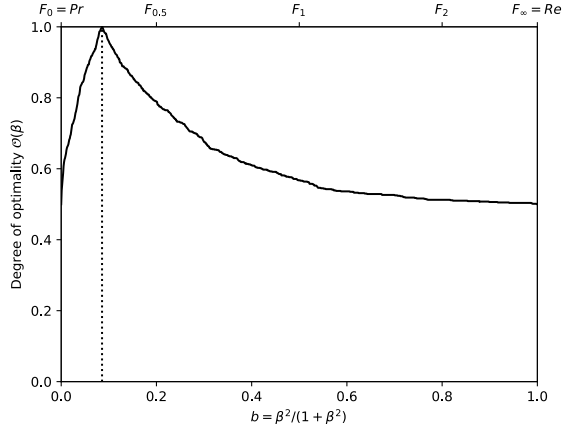
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

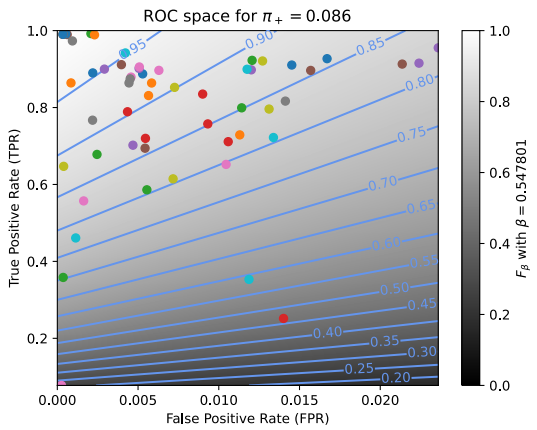


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

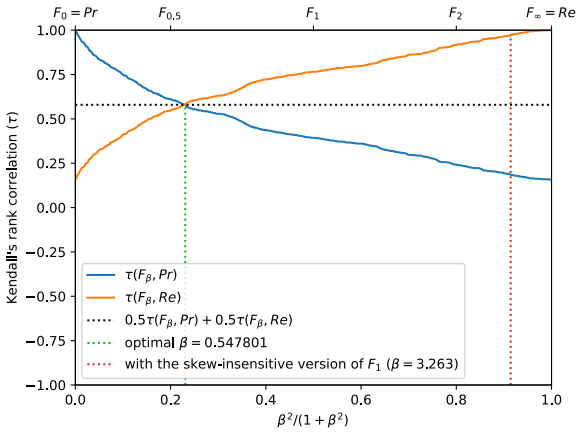


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

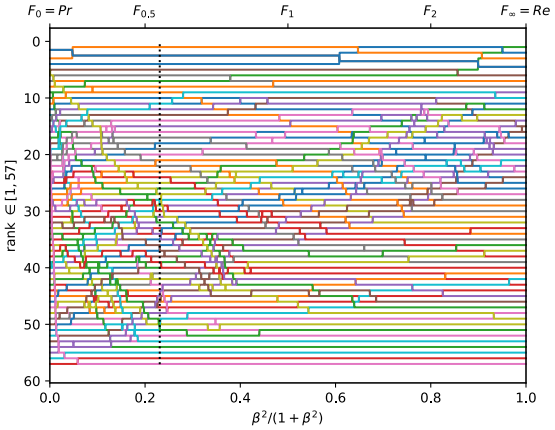
Figure A.3.36. Ranking of 57 BGS methods evaluated on the video "lakeSide" ($\pi_+ = 0.0192$).



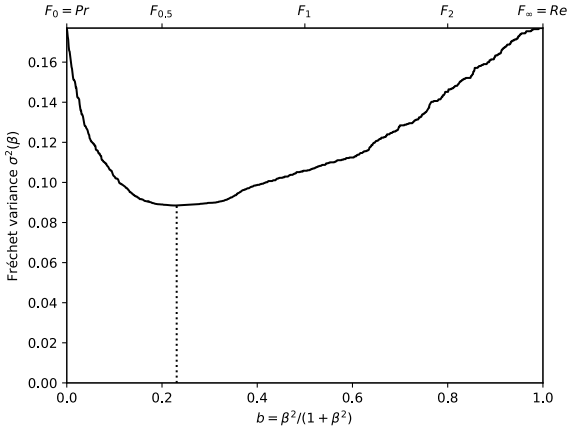
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



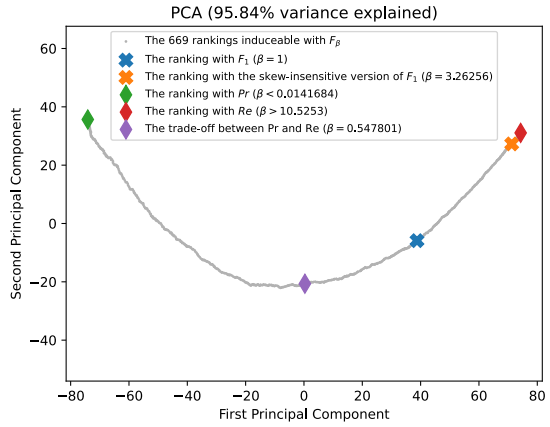
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



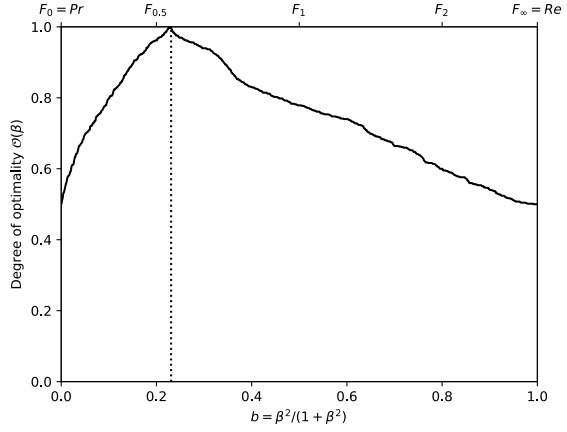
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

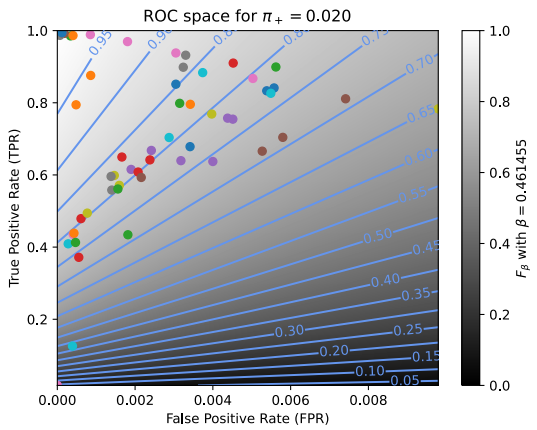


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

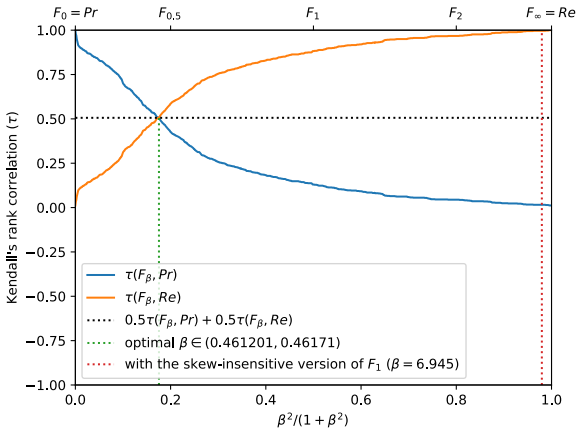


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

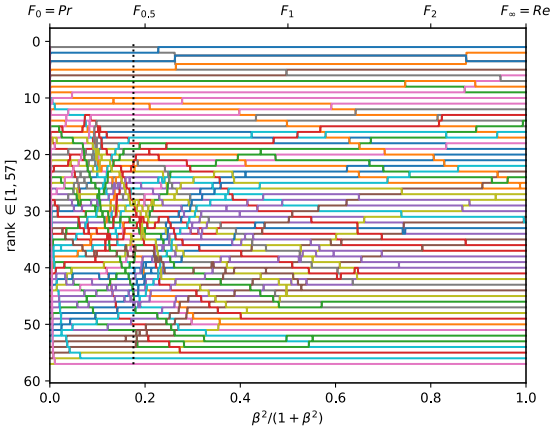
Figure A.3.37. Ranking of 57 BGS methods evaluated on the video "diningRoom" ($\pi_+ = 0.0859$).



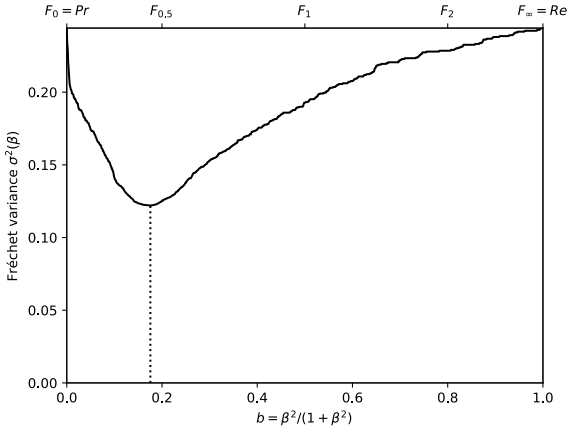
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



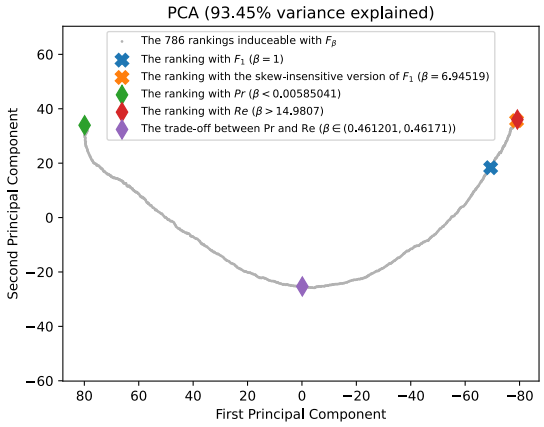
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



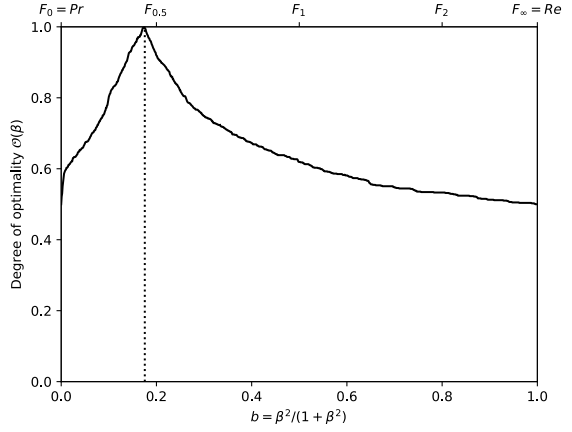
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

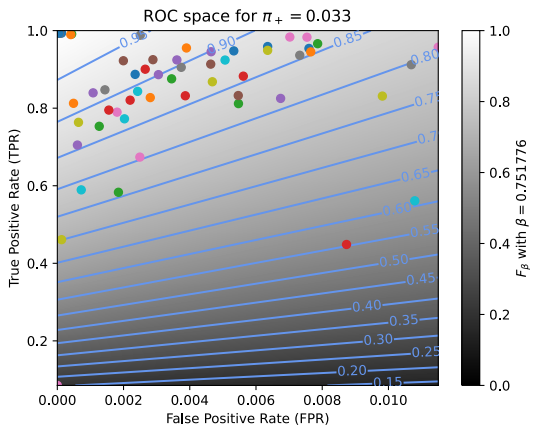


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

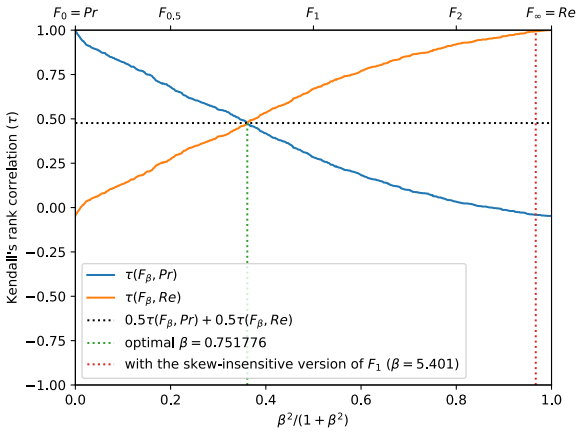


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

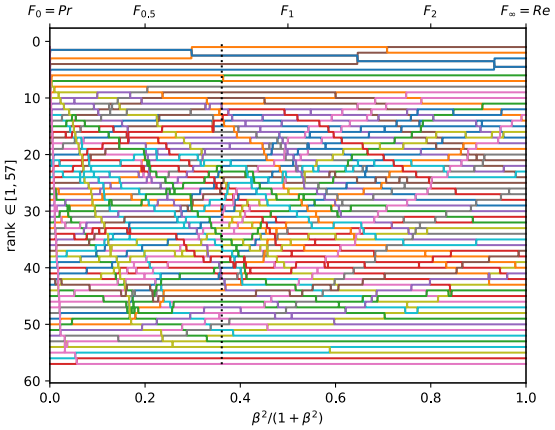
Figure A.3.38. Ranking of 57 BGS methods evaluated on the video "park" ($\pi_+ = 0.0203$).



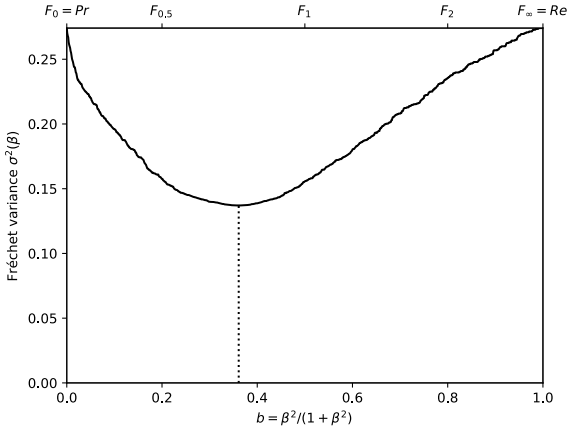
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



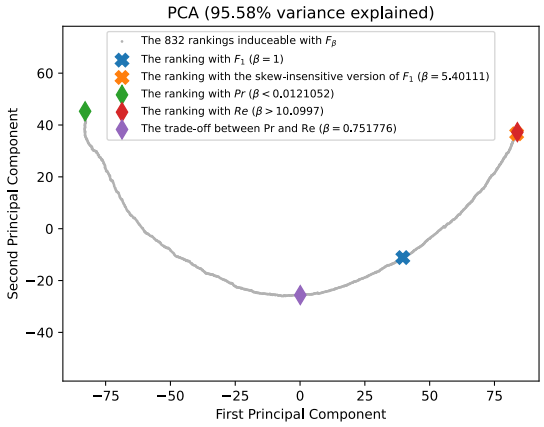
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



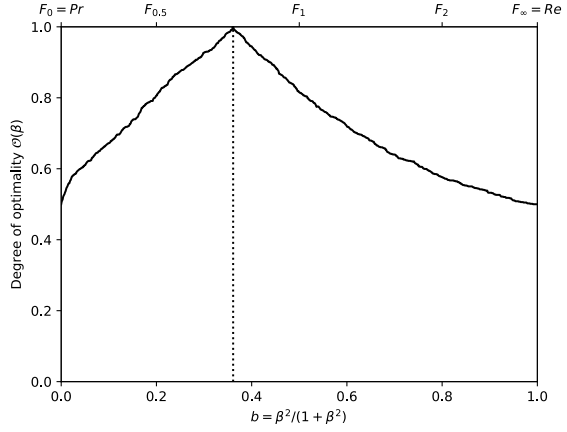
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

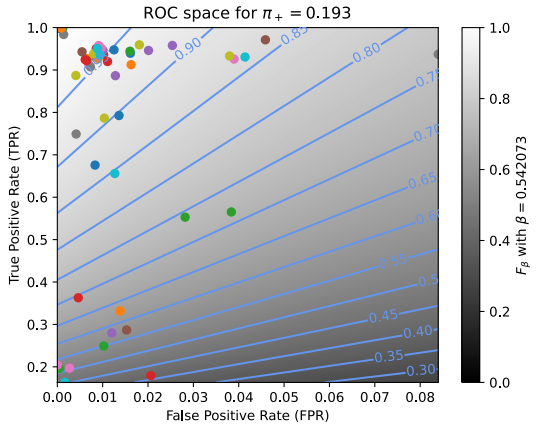


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

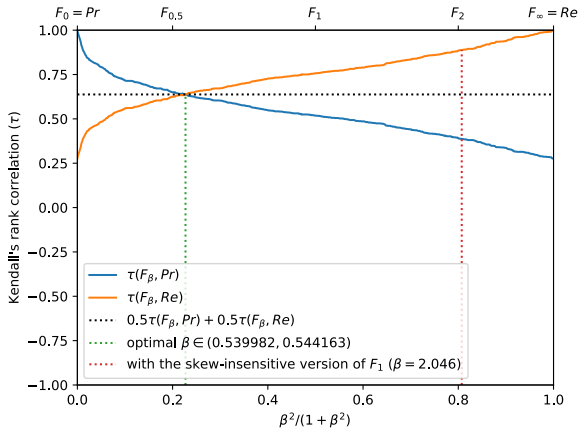


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

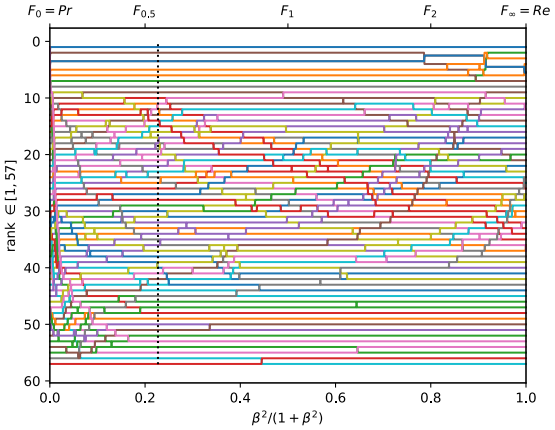
Figure A.3.39. Ranking of 57 BGS methods evaluated on the video "corridor" ($\pi_+ = 0.0331$).



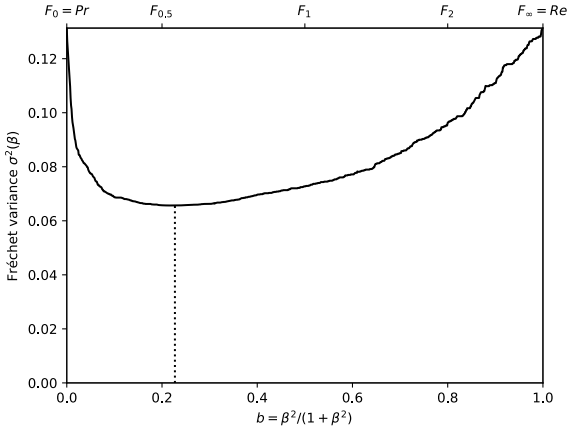
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



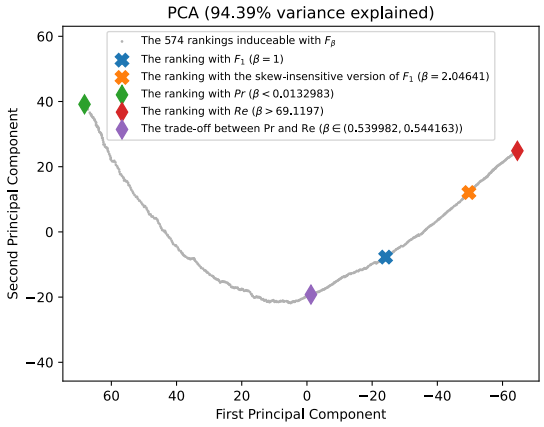
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



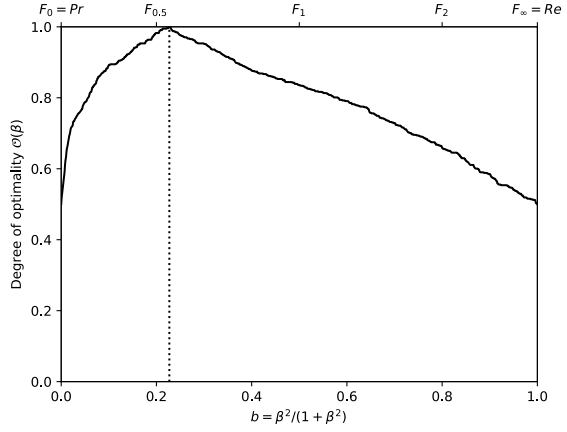
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

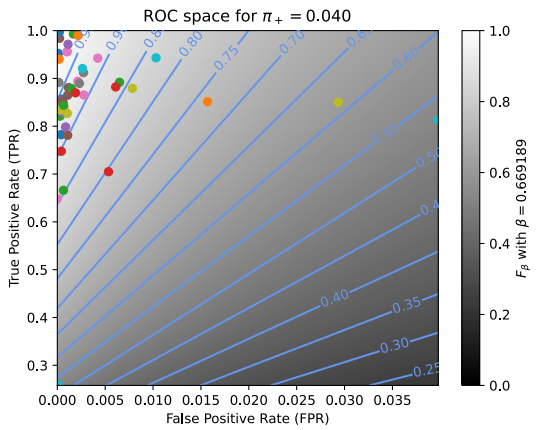


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

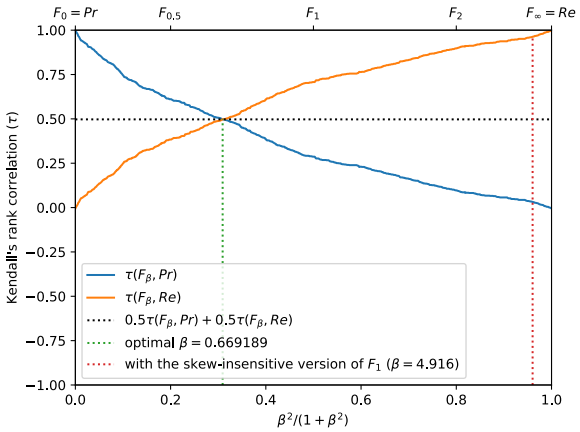


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

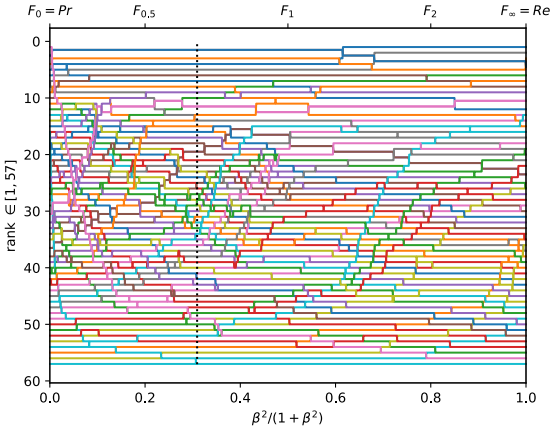
Figure A.3.40. Ranking of 57 BGS methods evaluated on the video "library" ($\pi_+ = 0.1928$).



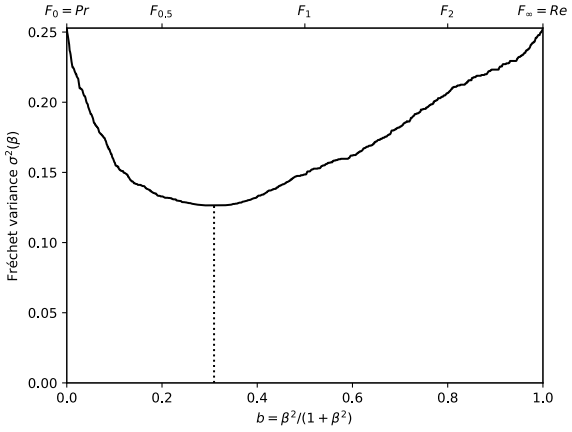
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



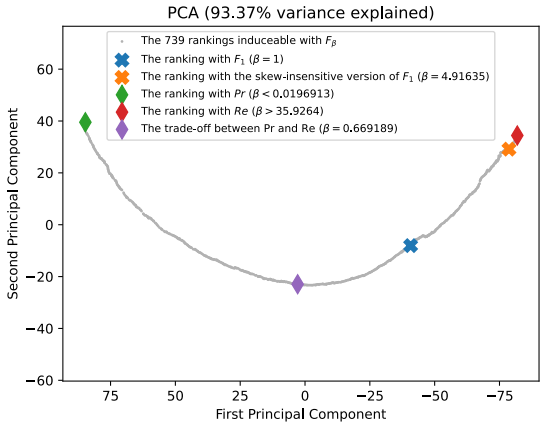
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



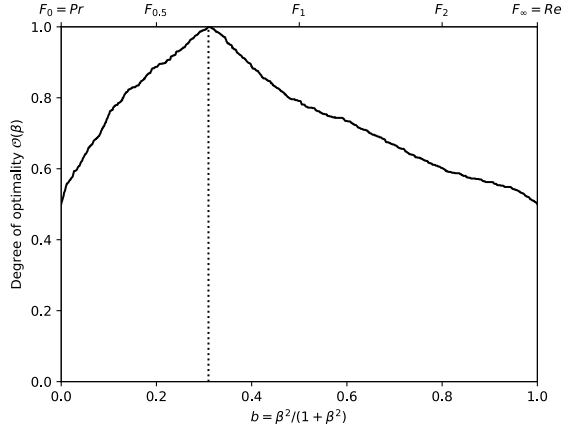
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

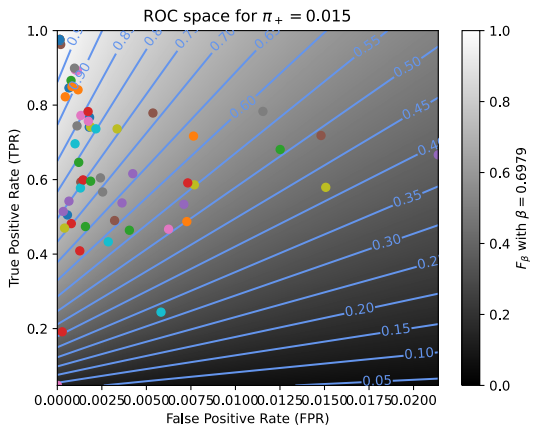


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

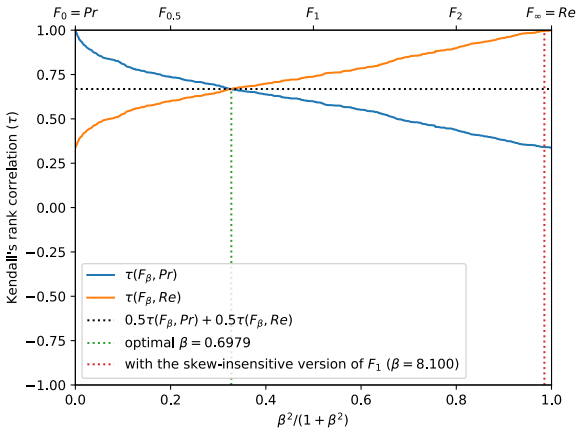


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

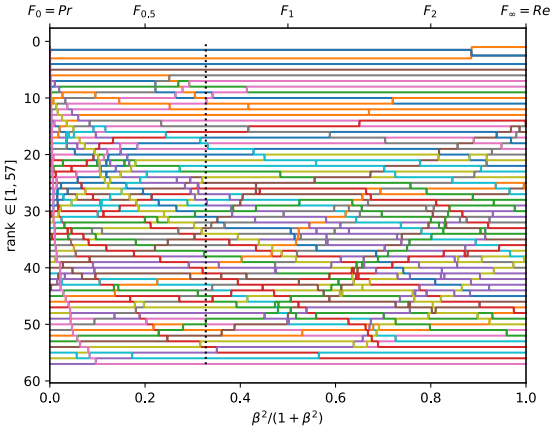
Figure A.3.41. Ranking of 57 BGS methods evaluated on the video "skating" ($\pi_+ = 0.0397$).



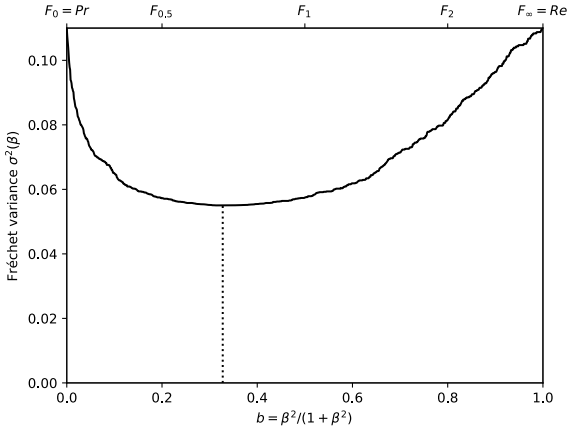
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



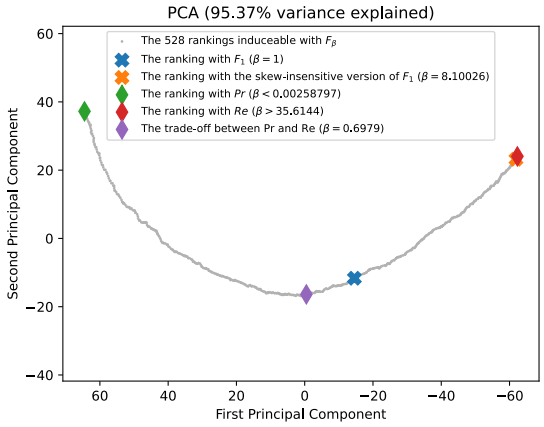
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



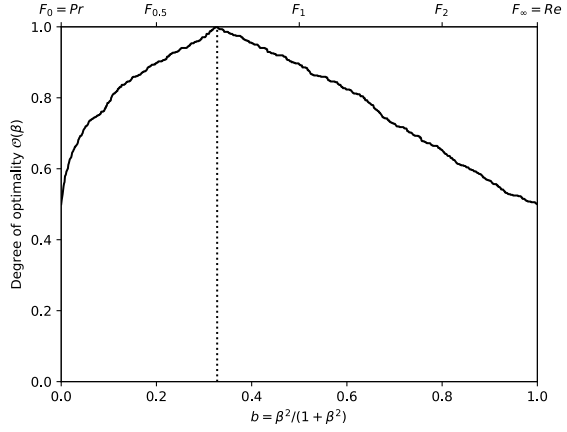
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

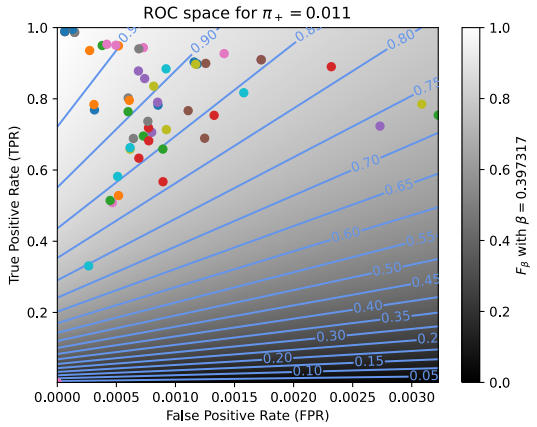


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

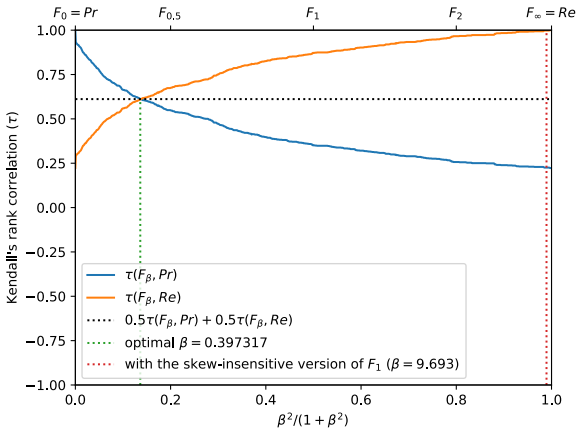


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

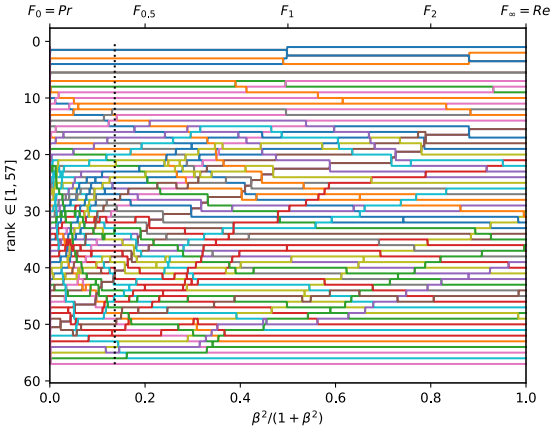
Figure A.3.42. Ranking of 57 BGS methods evaluated on the video "wetSnow" ($\pi_+ = 0.0150$).



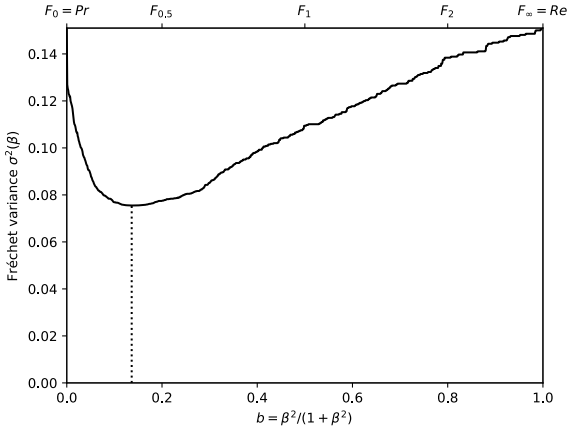
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



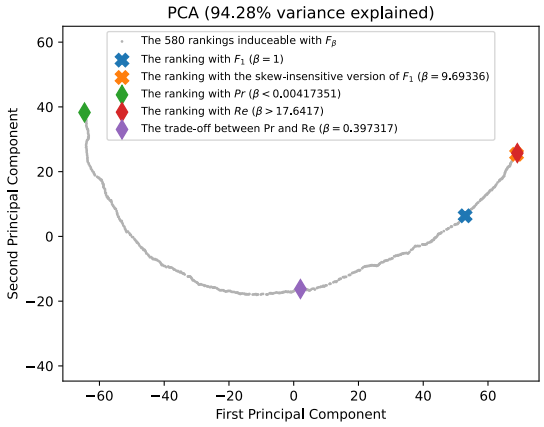
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



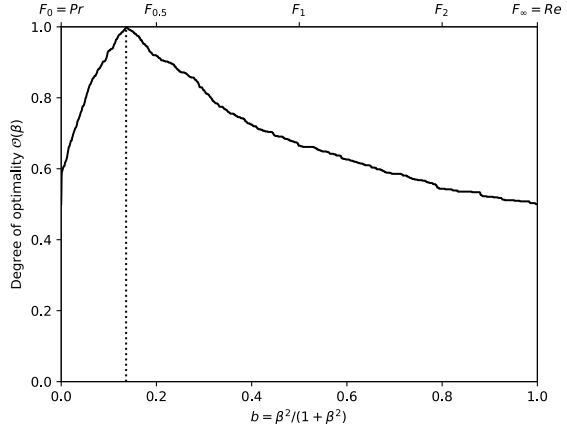
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

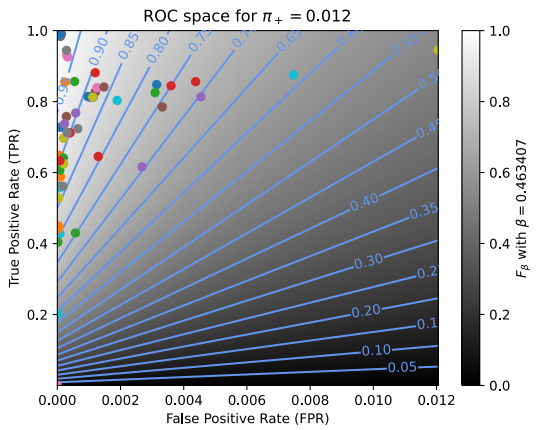


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

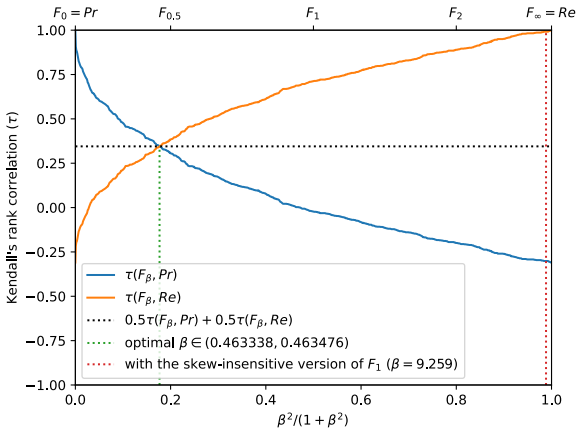


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

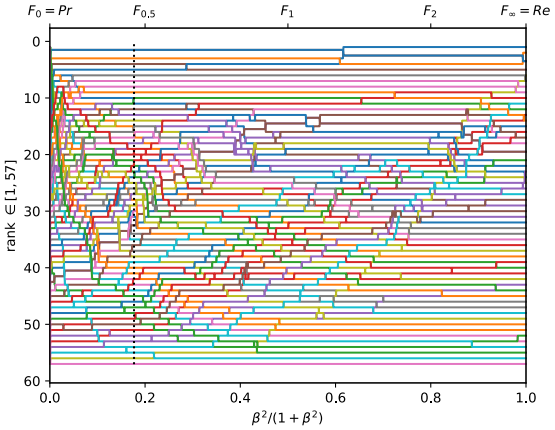
Figure A.3.43. Ranking of 57 BGS methods evaluated on the video "snowFall" ($\pi_+ = 0.0105$).



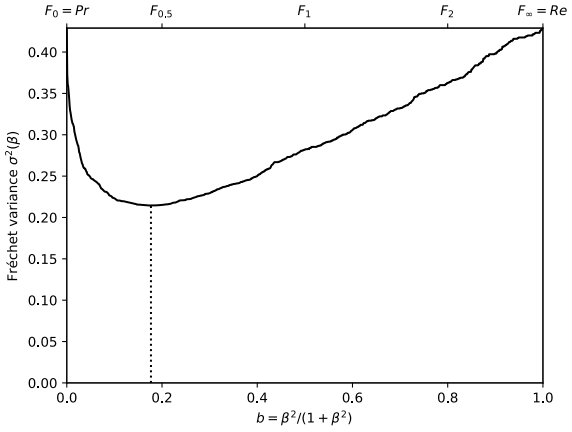
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



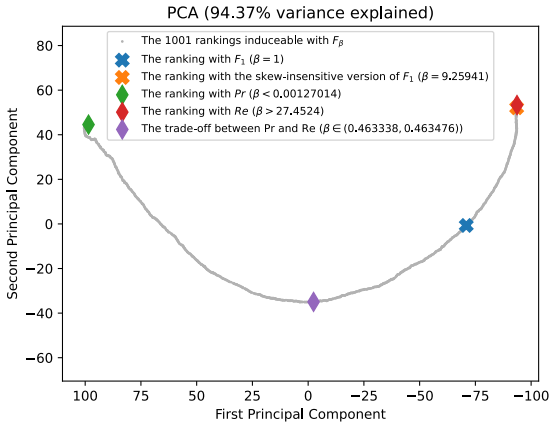
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



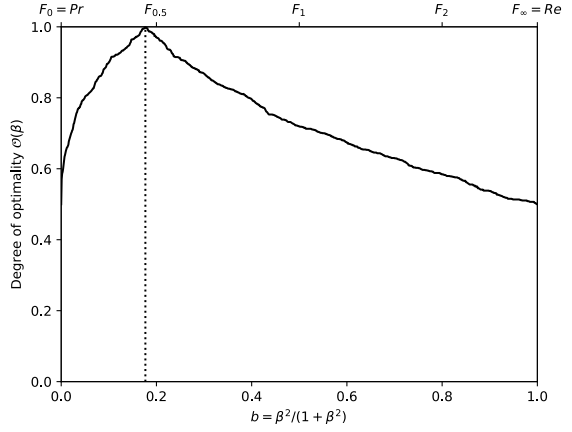
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

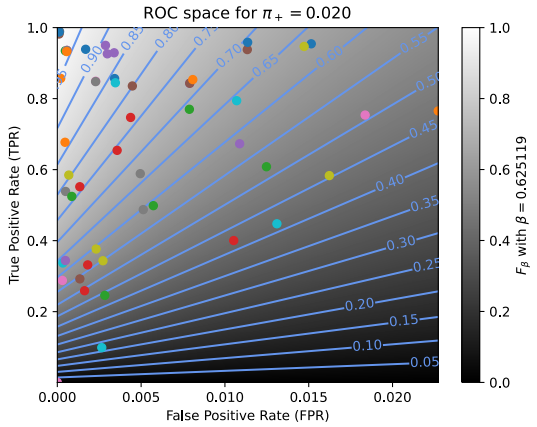


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

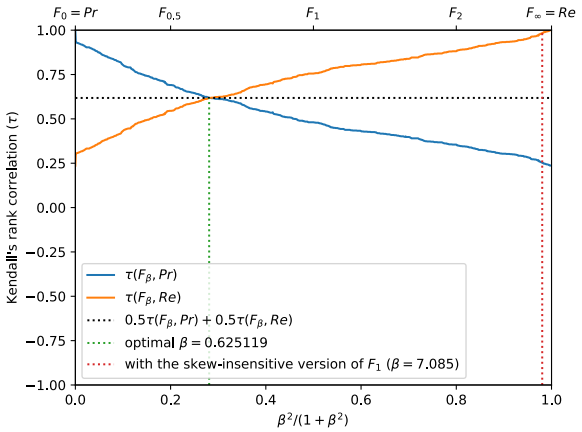


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

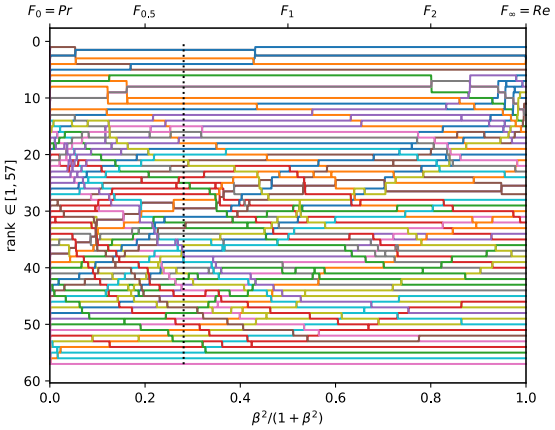
Figure A.3.44. Ranking of 57 BGS methods evaluated on the video "blizzard" ($\pi_+ = 0.0115$).



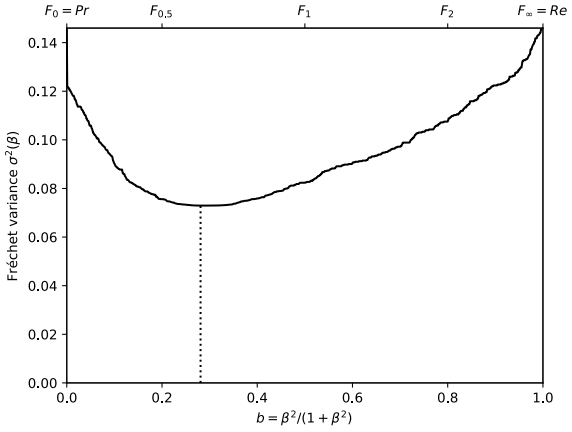
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



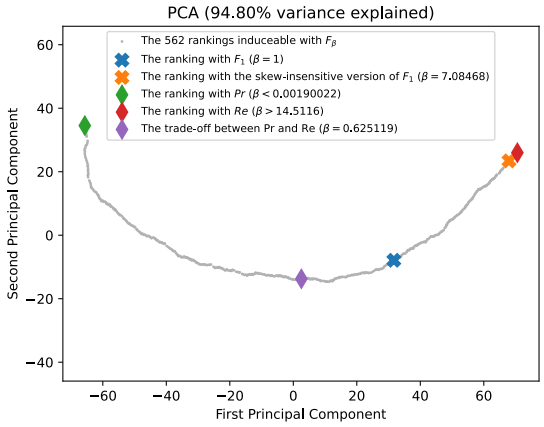
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



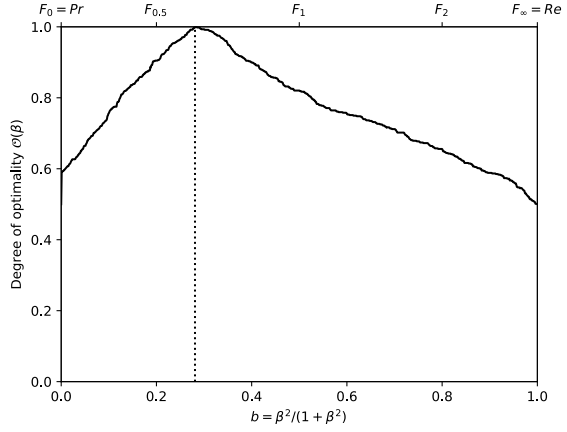
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

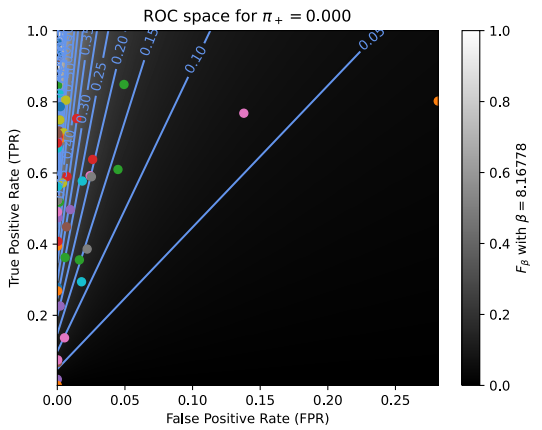


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

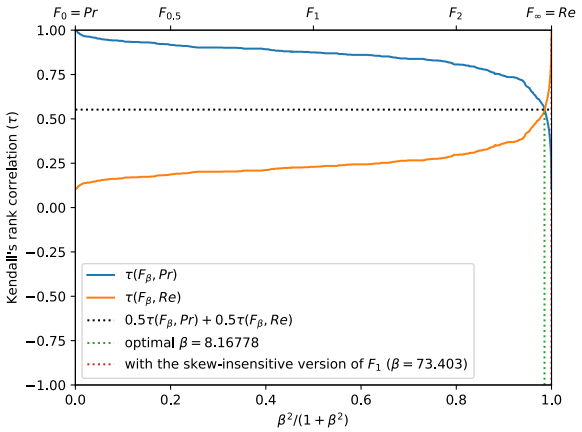


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

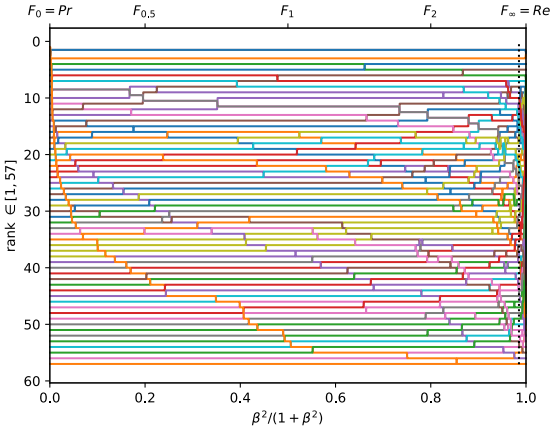
Figure A.3.45. Ranking of 57 BGS methods evaluated on the video "tunnelExit_0_35fps" ($\pi_+ = 0.0195$).



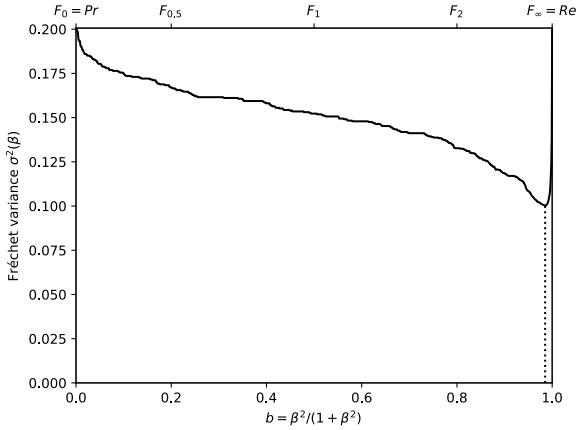
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



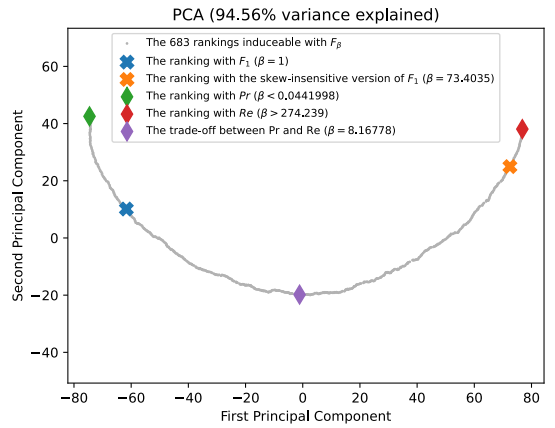
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



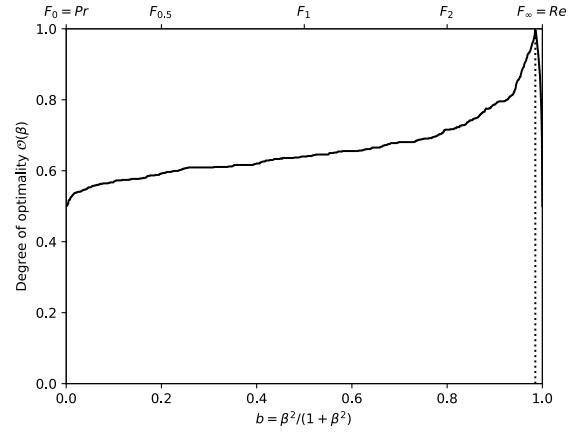
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

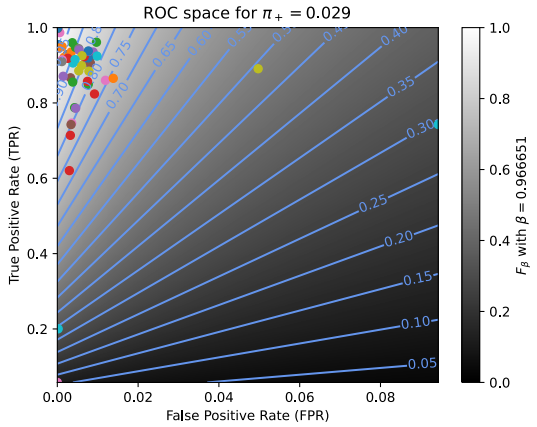


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

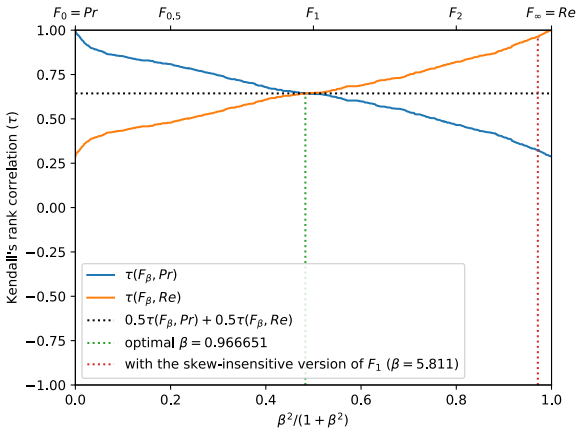


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

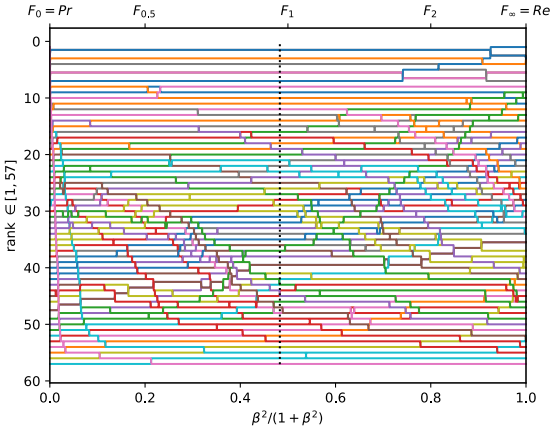
Figure A.3.46. Ranking of 57 BGS methods evaluated on the video "port_0_17fps" ($\pi_+ = 0.0002$).



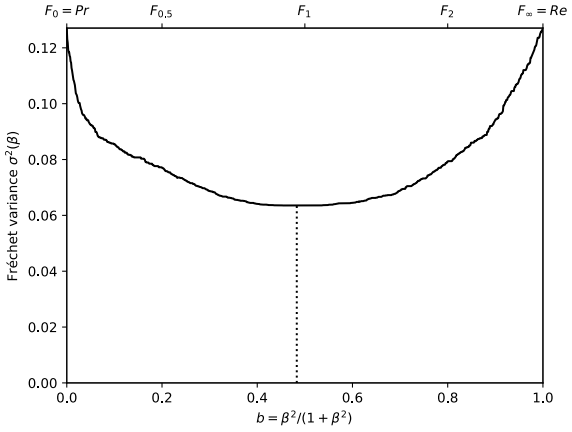
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



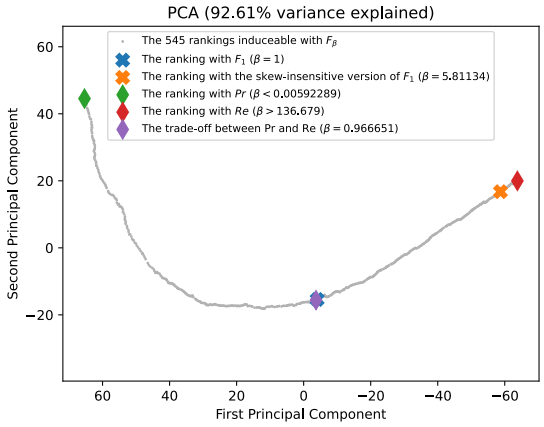
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



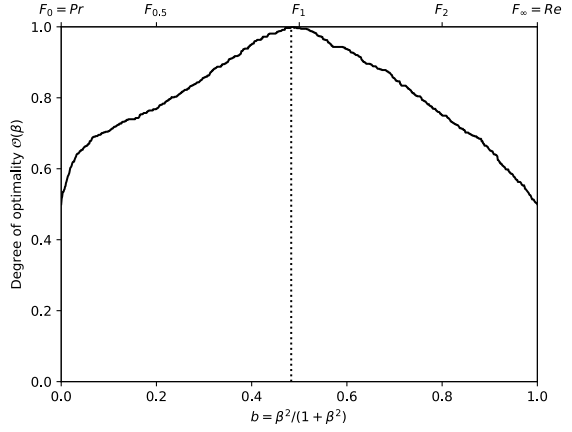
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

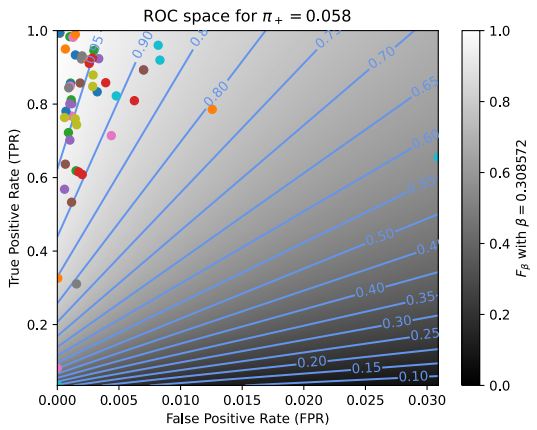


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

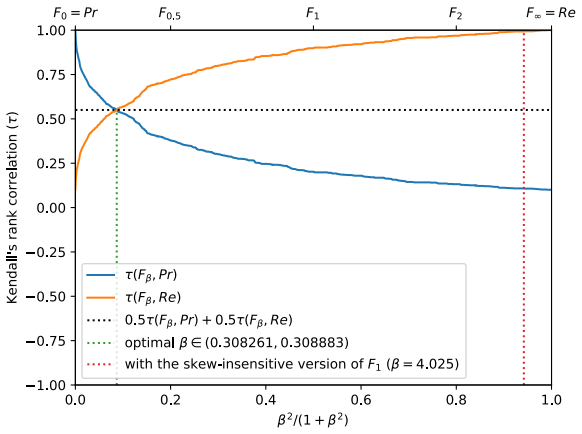


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

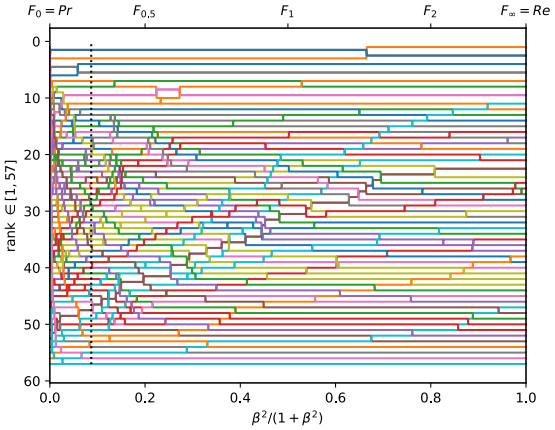
Figure A.3.47. Ranking of 57 BGS methods evaluated on the video "tramCrossroad_1fps" ($\pi_+ = 0.0288$).



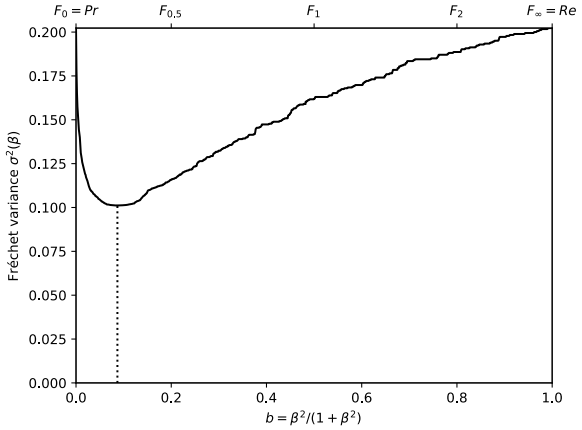
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



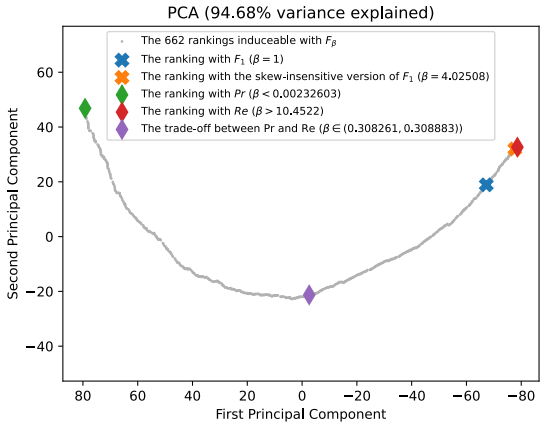
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



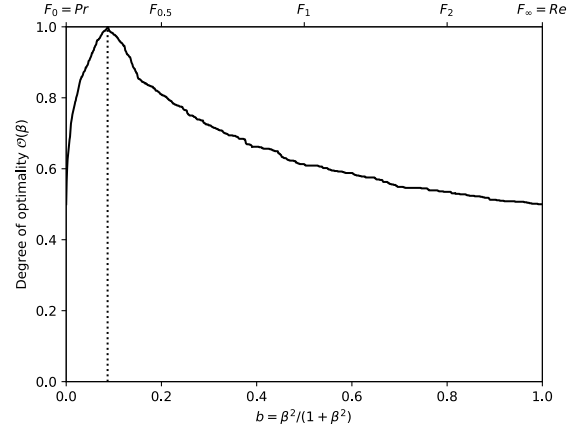
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

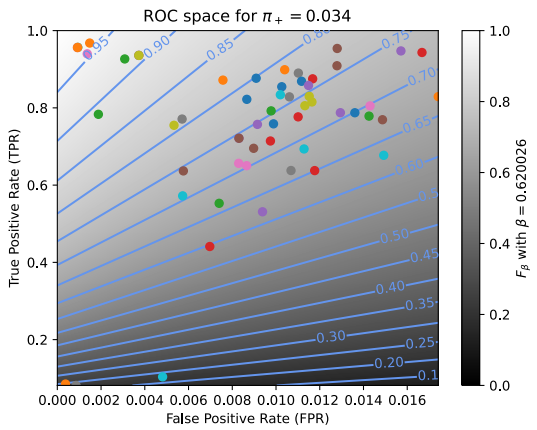


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

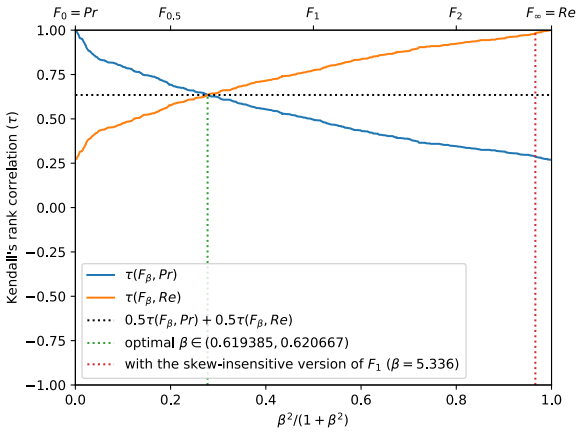


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

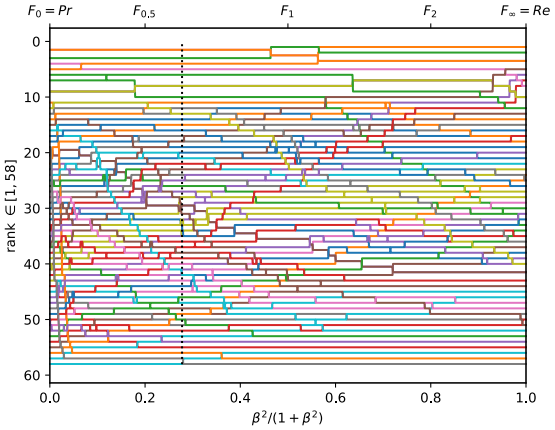
Figure A.3.48. Ranking of 57 BGS methods evaluated on the video "turnpike_0_5fps" ($\pi_+ = 0.0581$).



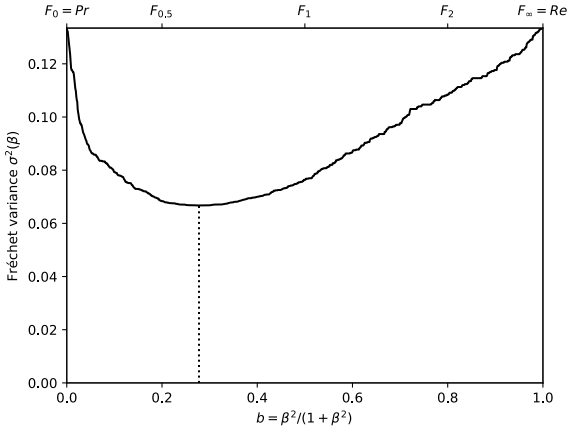
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



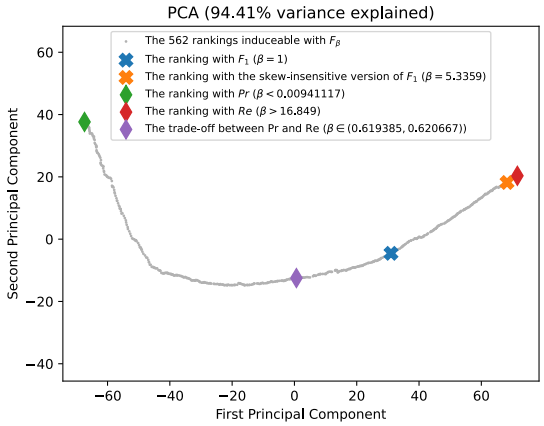
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



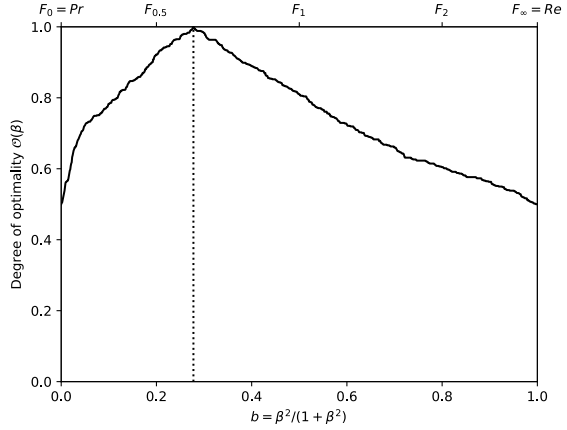
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

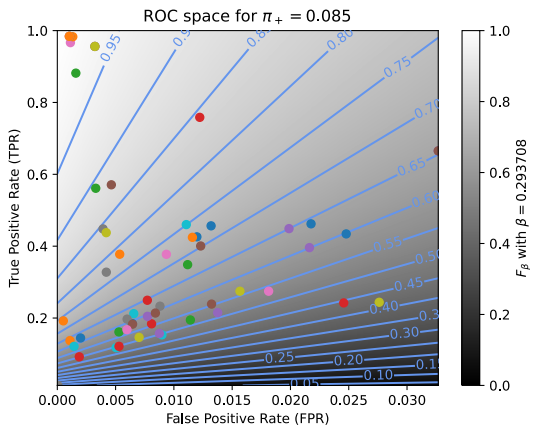


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

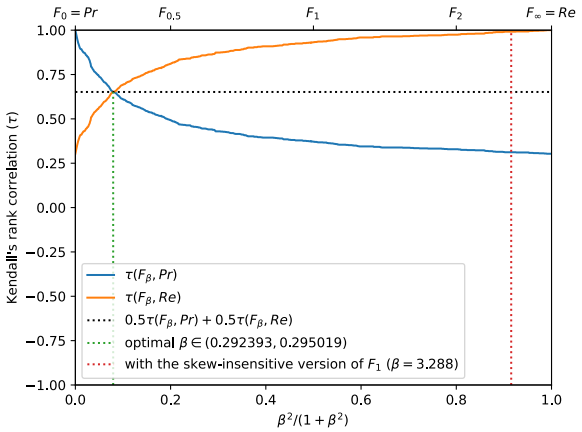


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

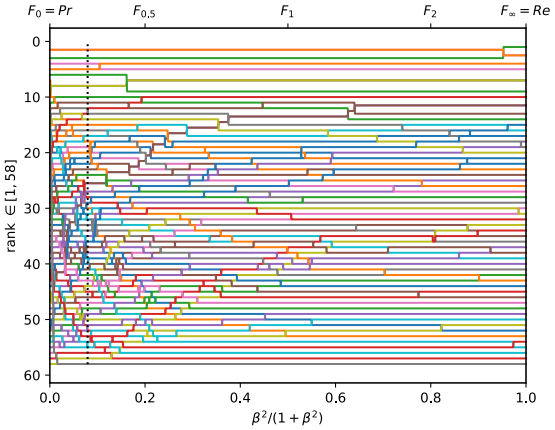
Figure A.3.49. Ranking of 58 BGS methods evaluated on the video "tramStation" ($\pi_+ = 0.0339$).



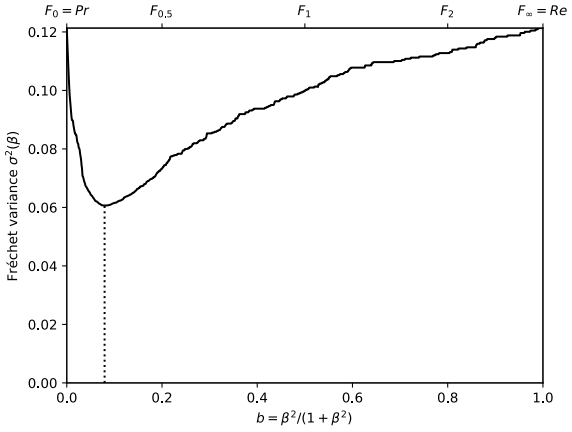
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



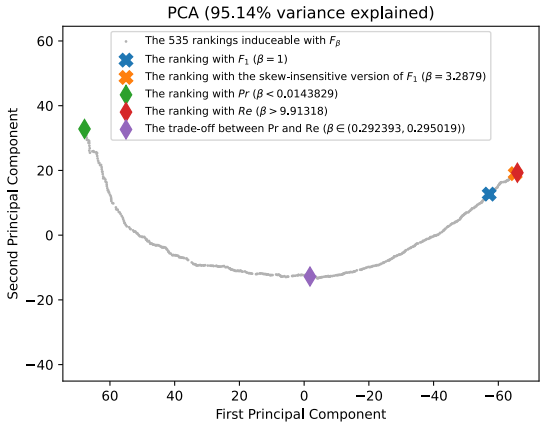
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



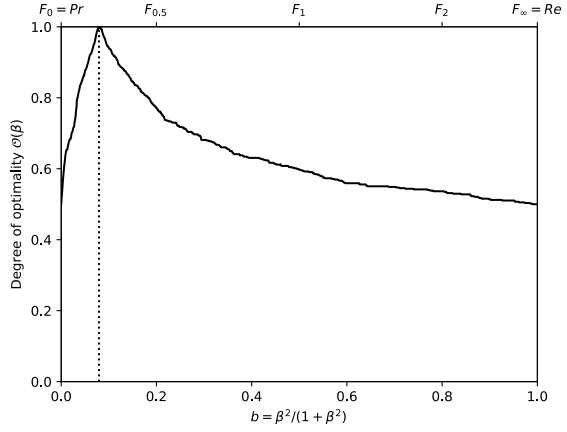
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

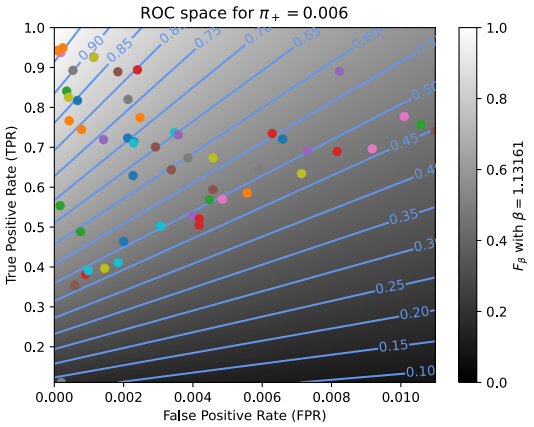


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

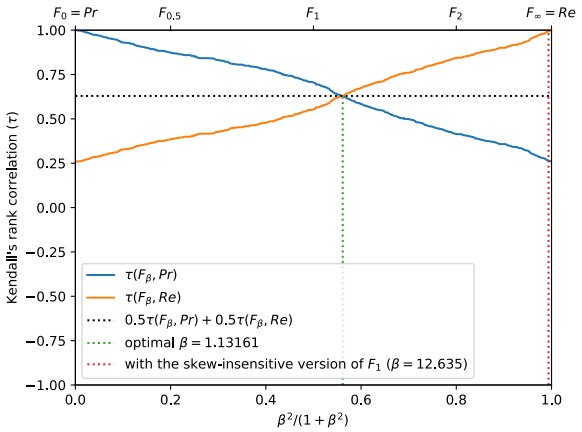


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

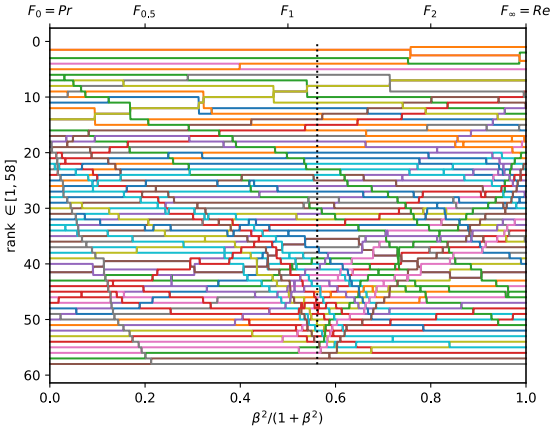
Figure A.3.50. Ranking of 58 BGS methods evaluated on the video "busyBoulvard" ($\pi_+ = 0.0847$).



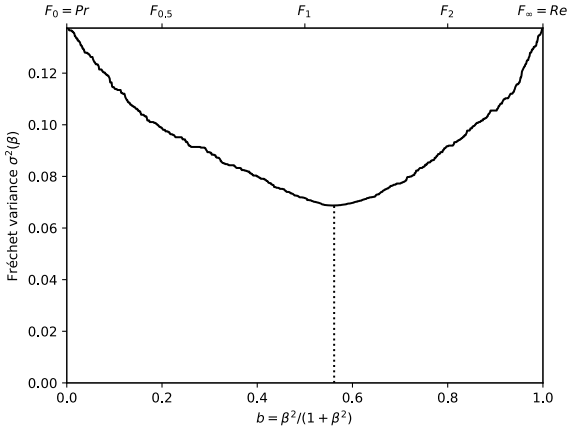
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



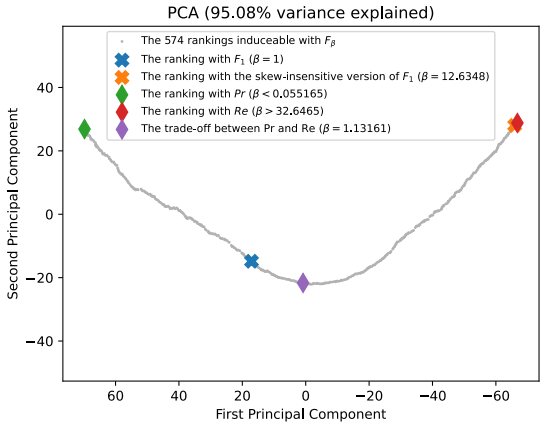
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



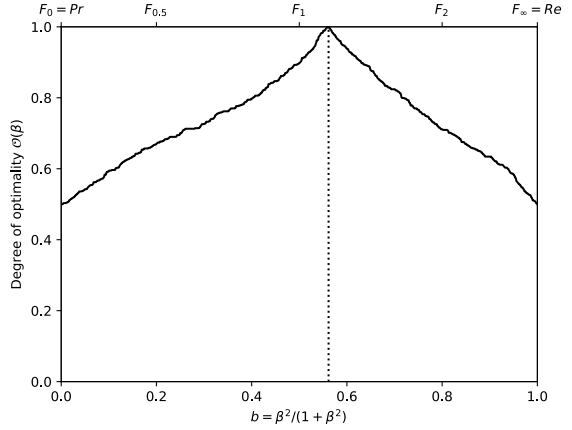
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

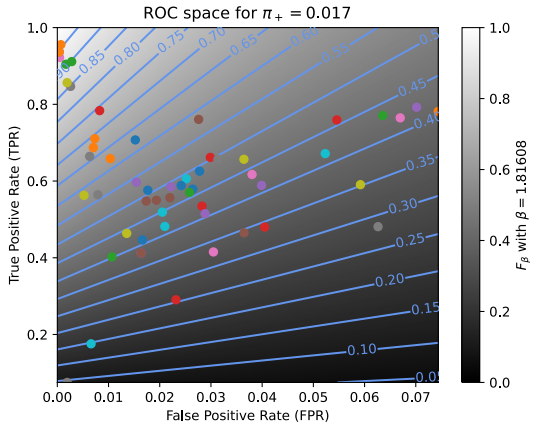


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

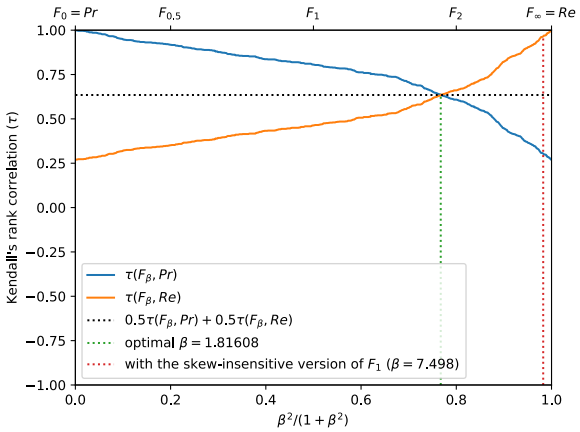


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

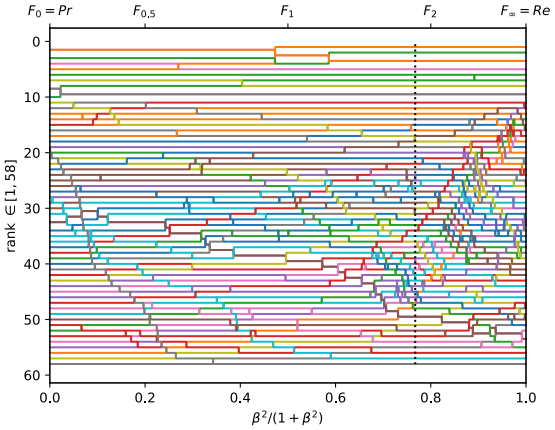
Figure A.3.51. Ranking of 58 BGS methods evaluated on the video "streetCornerAtNight" ($\pi_+ = 0.0062$).



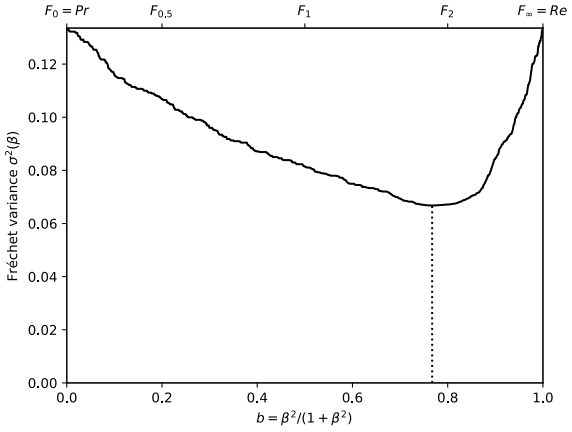
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



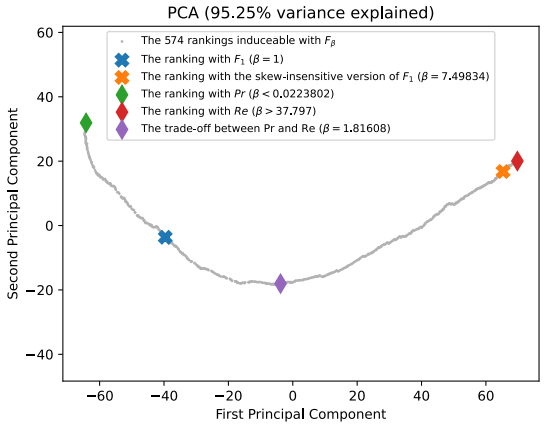
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



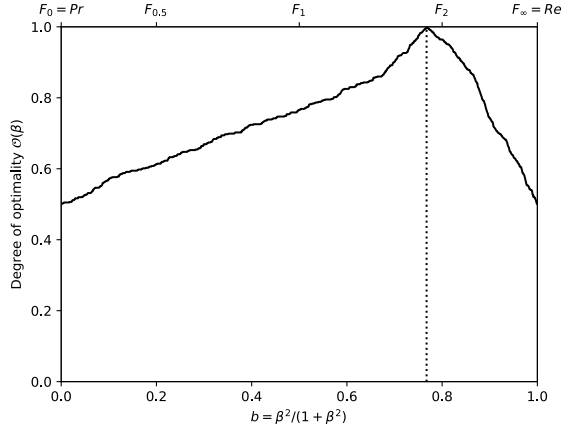
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

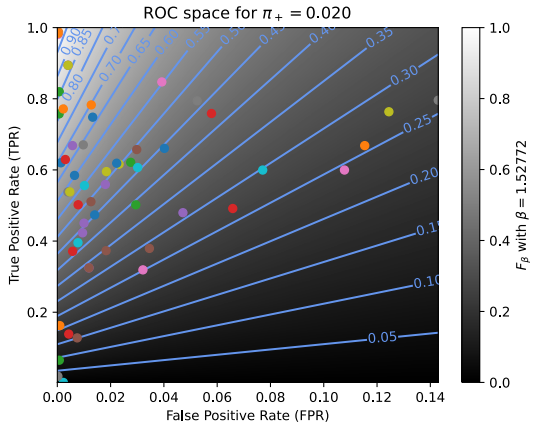


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

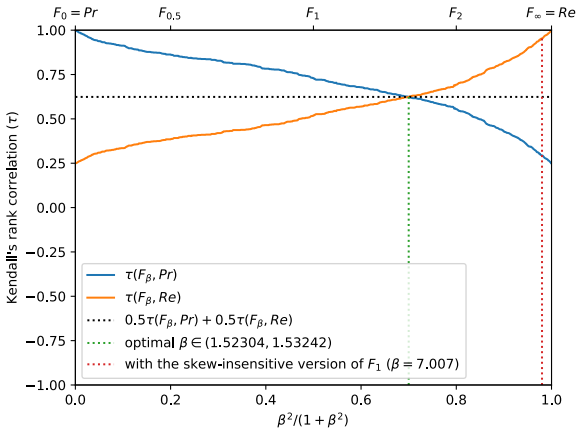


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

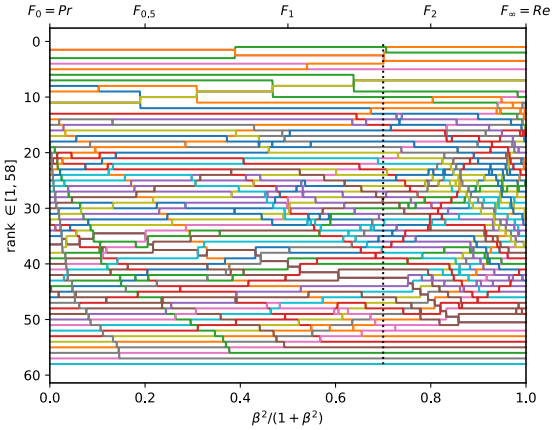
Figure A.3.52. Ranking of 58 BGS methods evaluated on the video "fluidHighway" ($\pi_+ = 0.0175$).



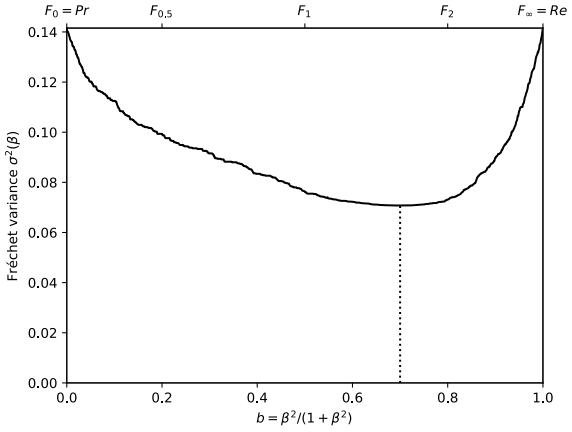
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



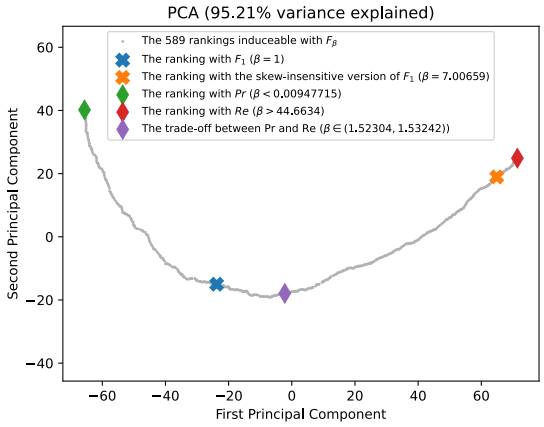
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



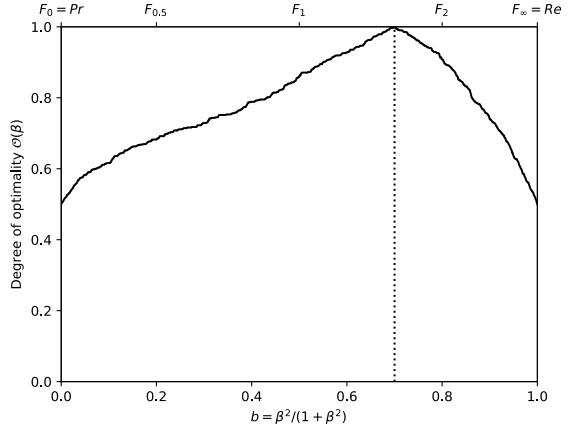
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

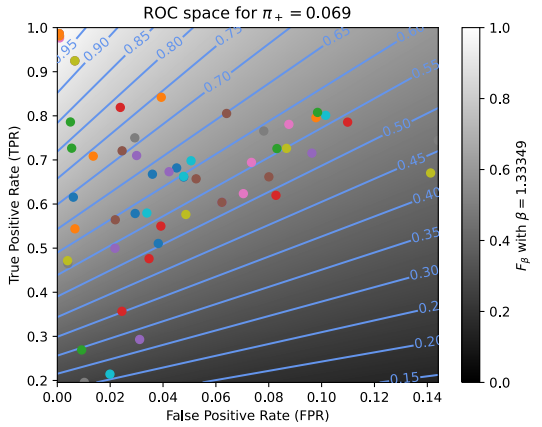


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

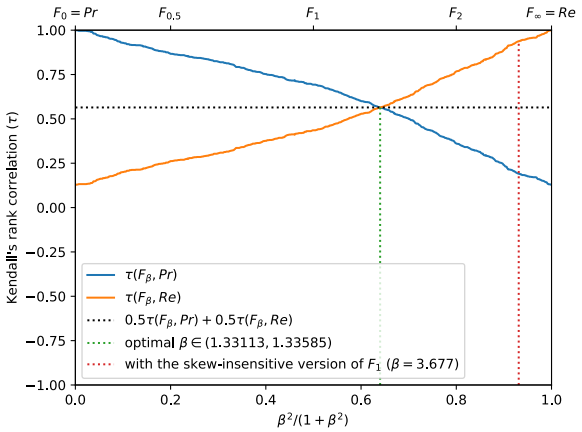


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

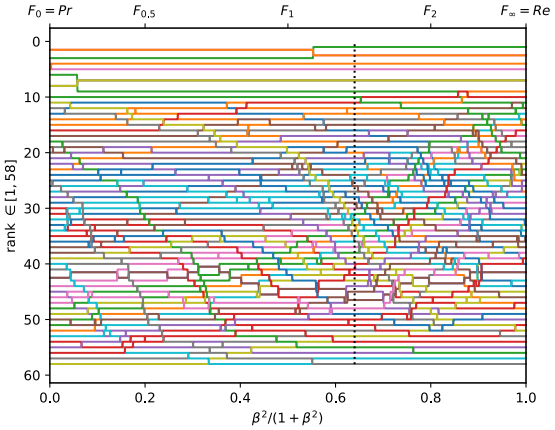
Figure A.3.53. Ranking of 58 BGS methods evaluated on the video "bridgeEntry" ($\pi_+ = 0.0200$).



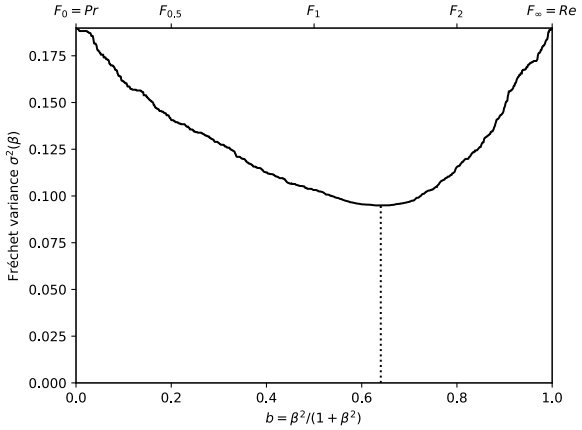
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



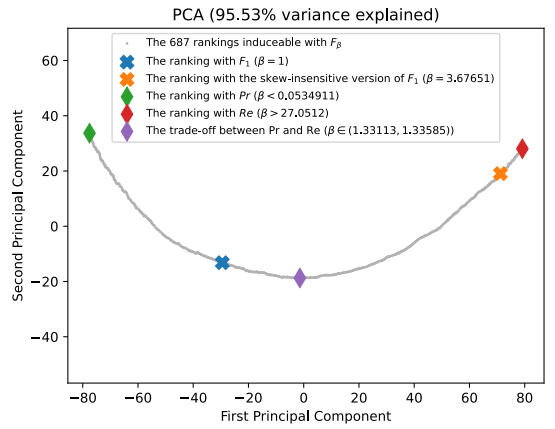
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



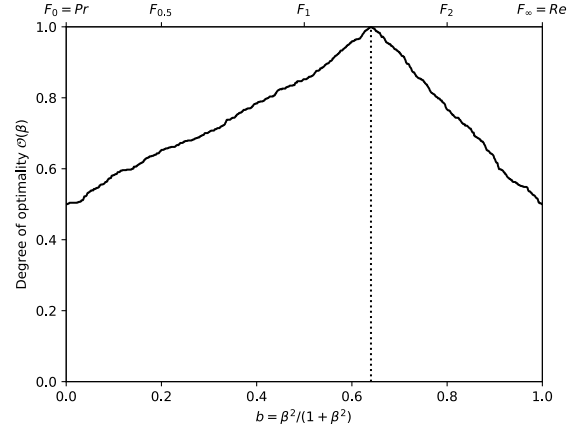
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

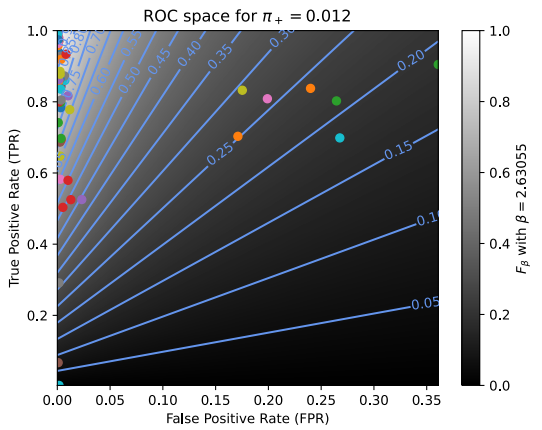


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

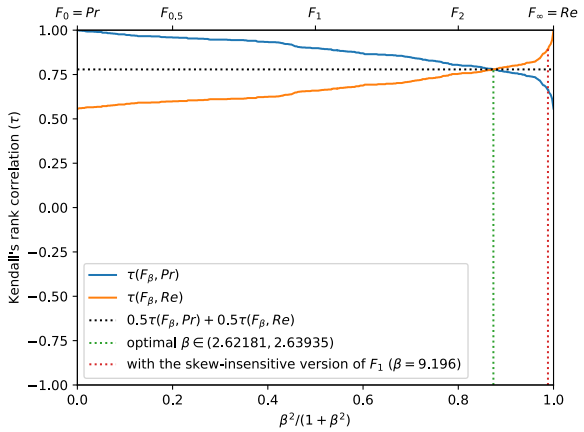


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

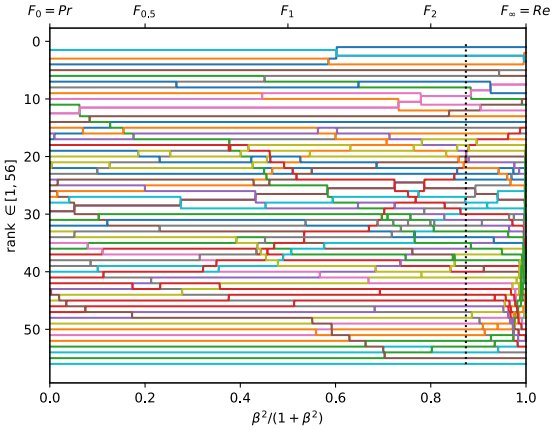
Figure A.3.54. Ranking of 58 BGS methods evaluated on the video "winterStreet" ($\pi_+ = 0.0689$).



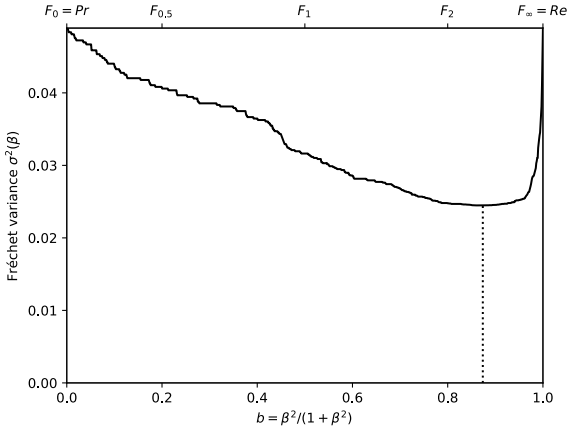
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



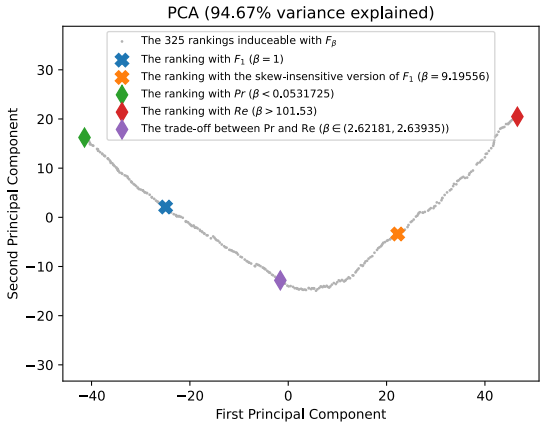
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



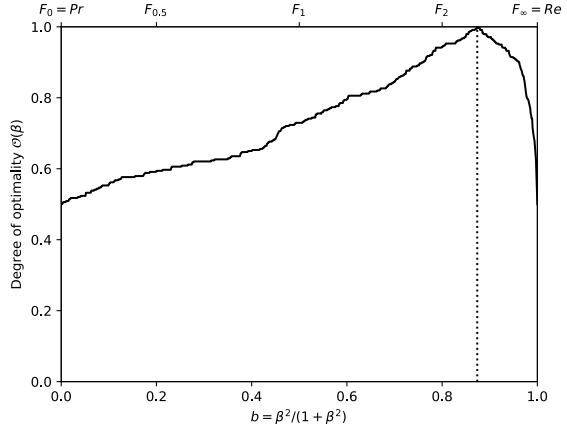
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

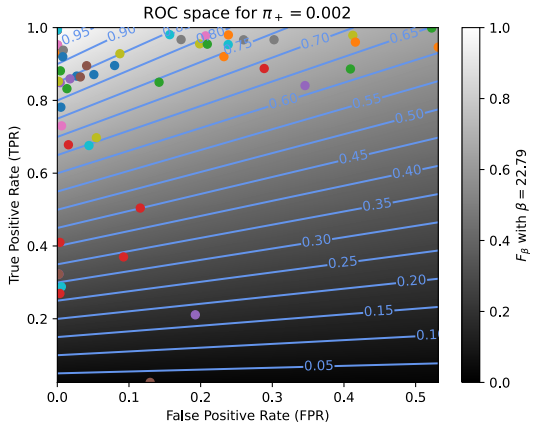


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

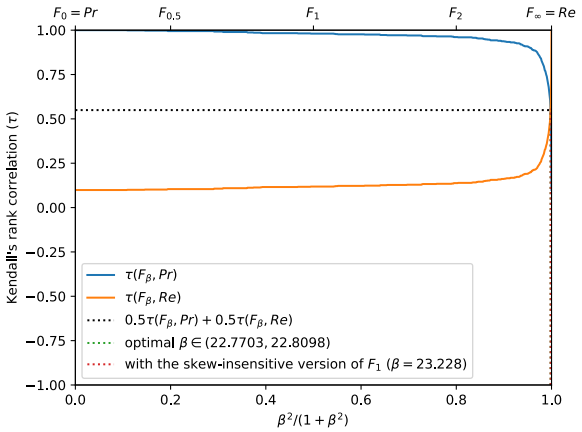


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

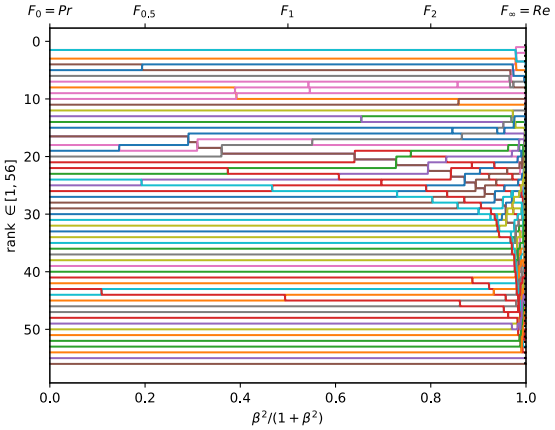
Figure A.3.55. Ranking of 56 BGS methods evaluated on the video "twoPositionPTZCam" ($\pi_+ = 0.0117$).



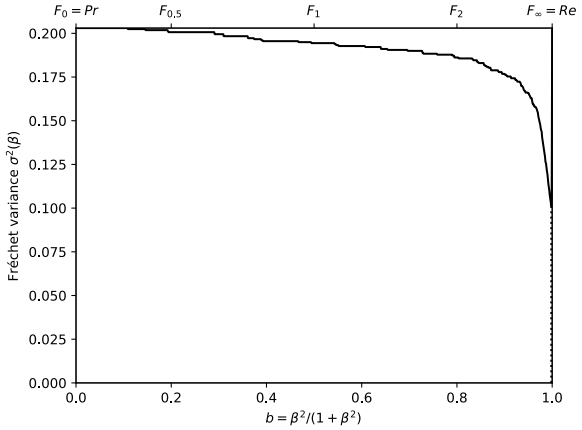
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



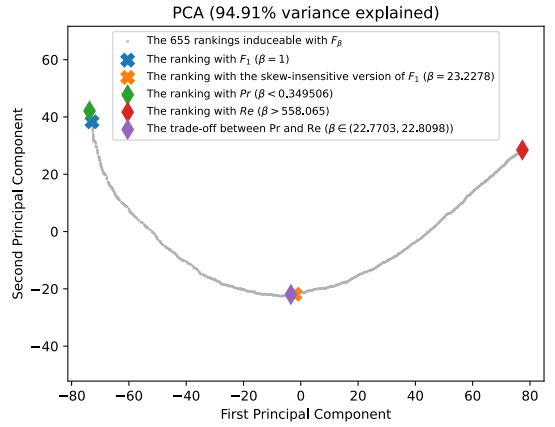
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



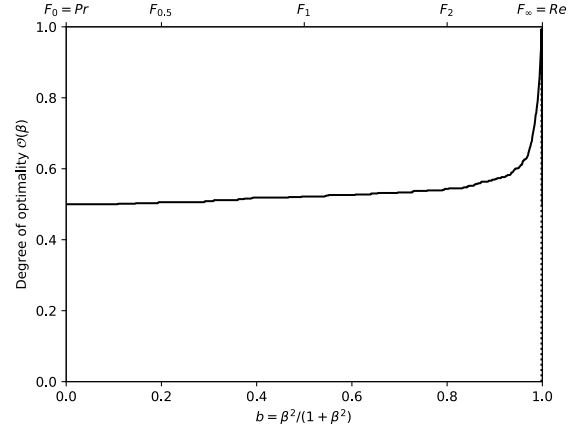
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

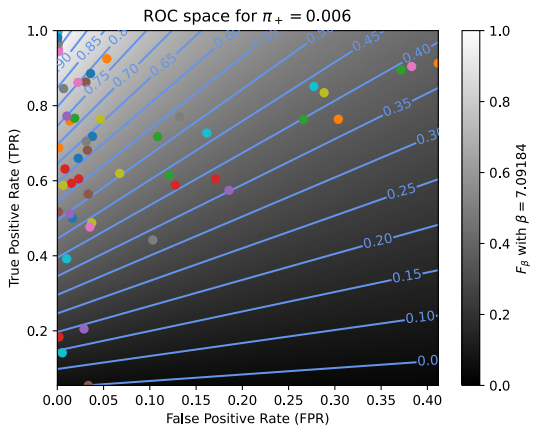


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

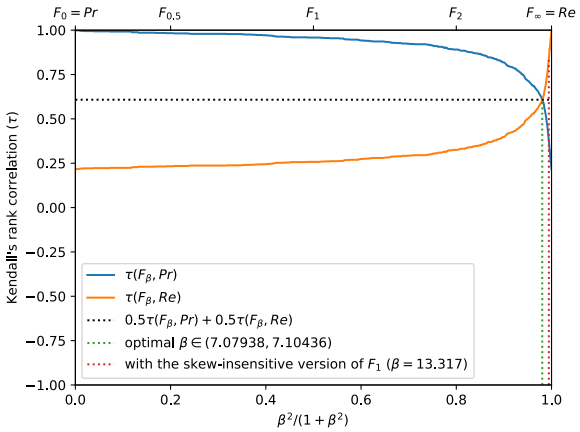


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

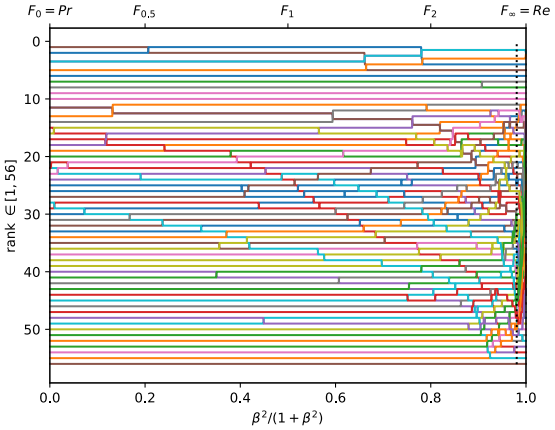
Figure A.3.56. Ranking of 56 BGS methods evaluated on the video "zoomInZoomOut" ($\pi_+ = 0.0019$).



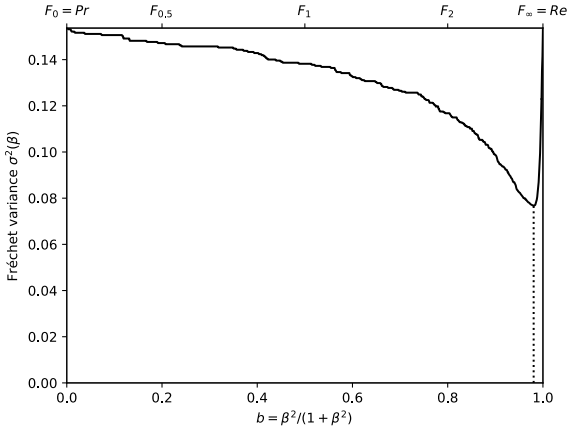
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



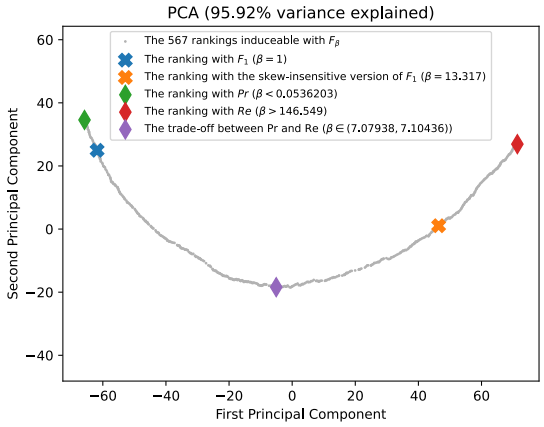
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



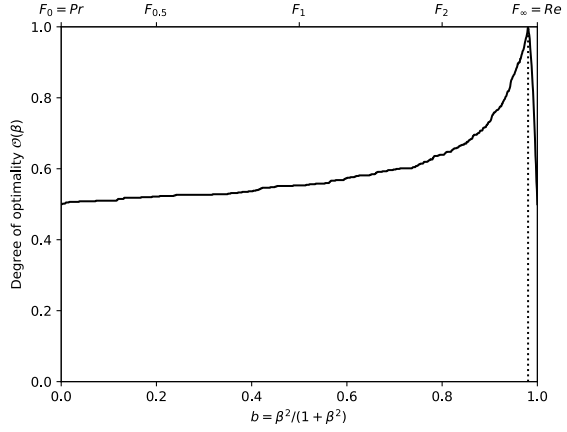
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

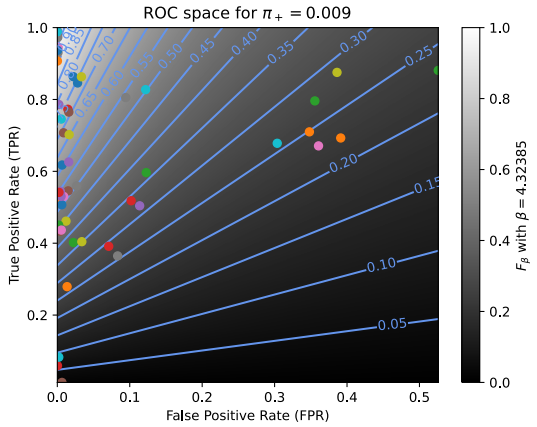


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

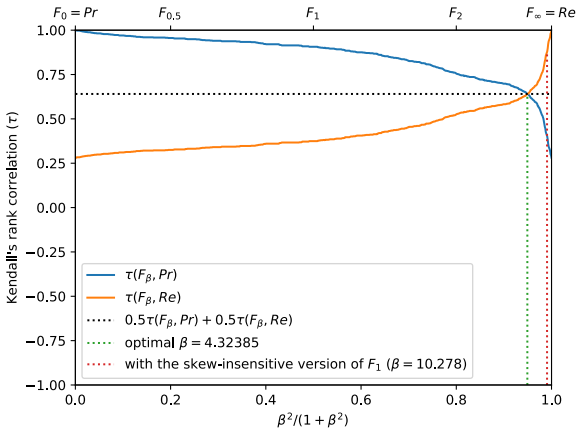


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

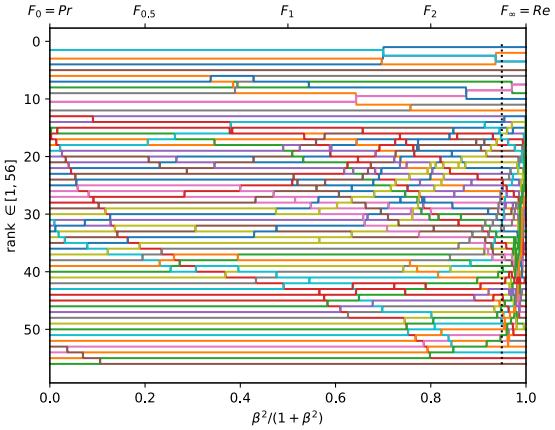
Figure A.3.57. Ranking of 56 BGS methods evaluated on the video "continuousPan" ($\pi_+ = 0.0056$).



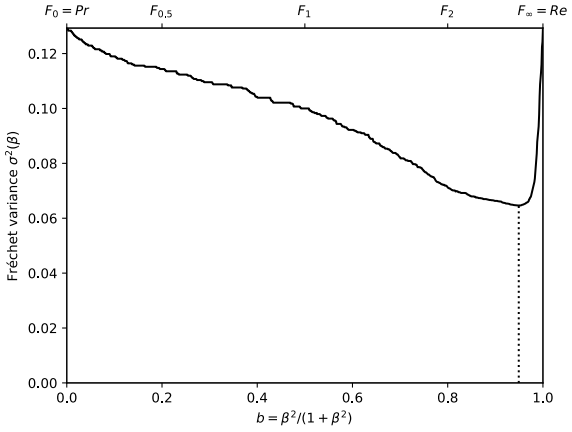
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



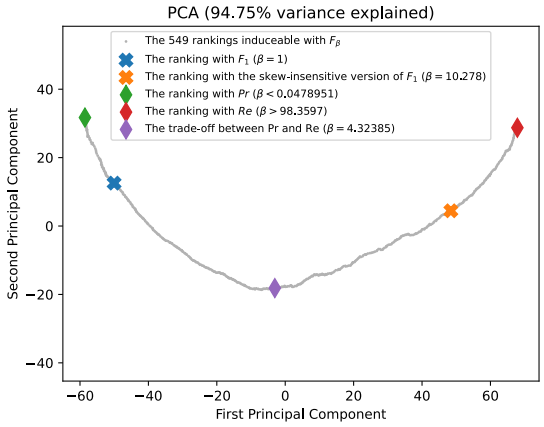
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



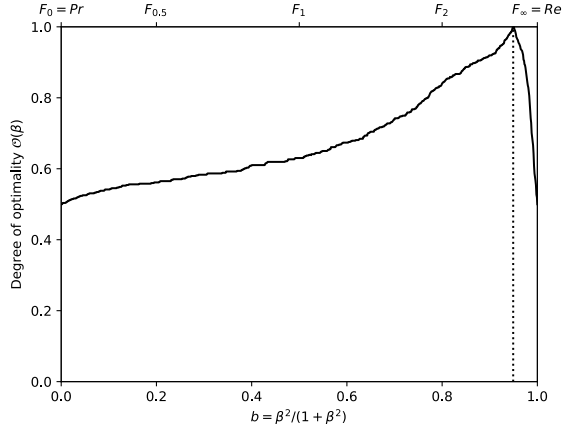
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

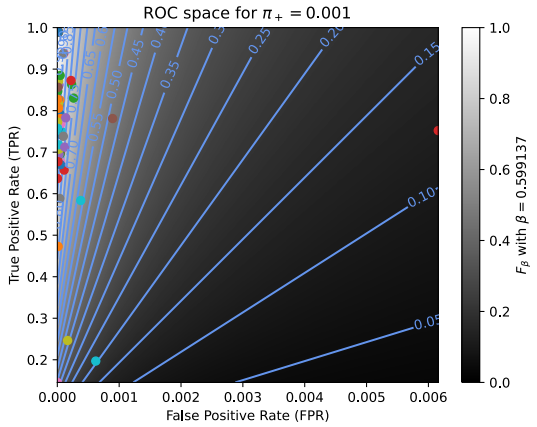


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

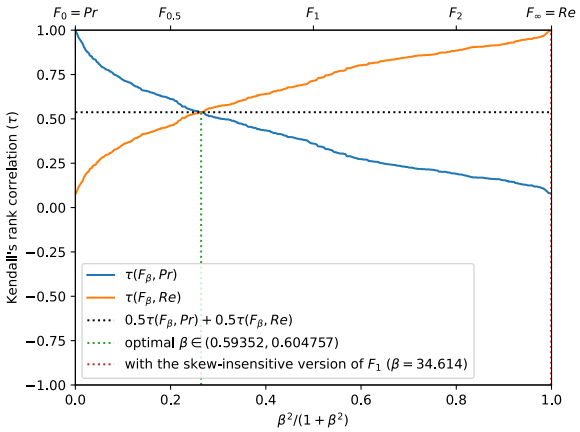


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

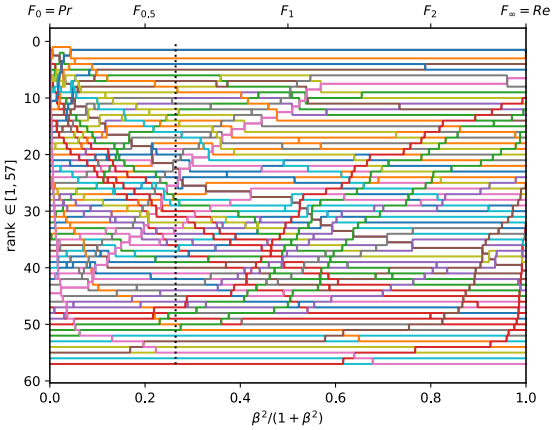
Figure A.3.58. Ranking of 56 BGS methods evaluated on the video "intermittentPan" ($\pi_+ = 0.0094$).



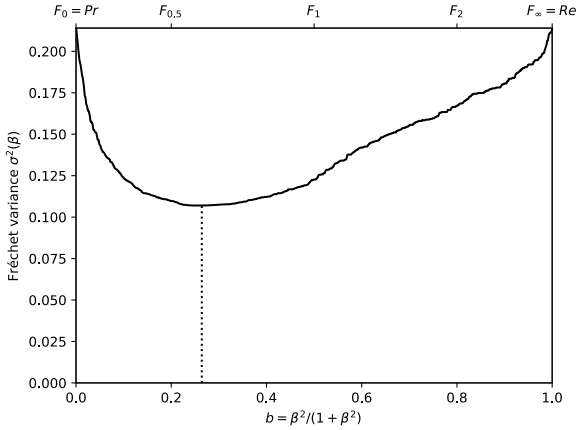
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



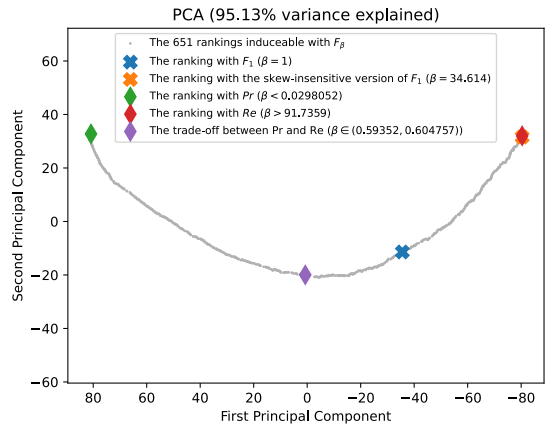
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



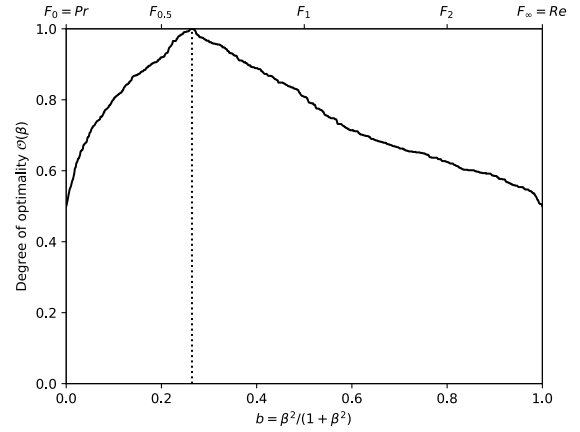
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

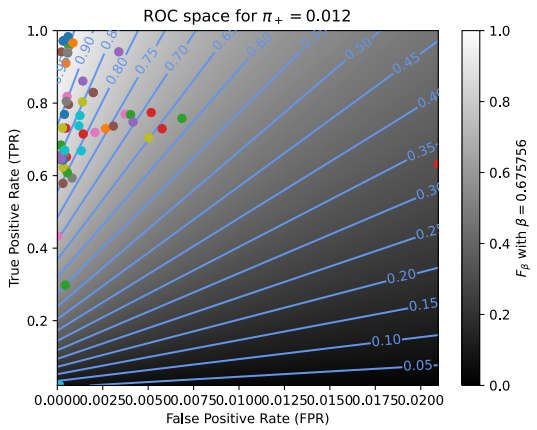


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

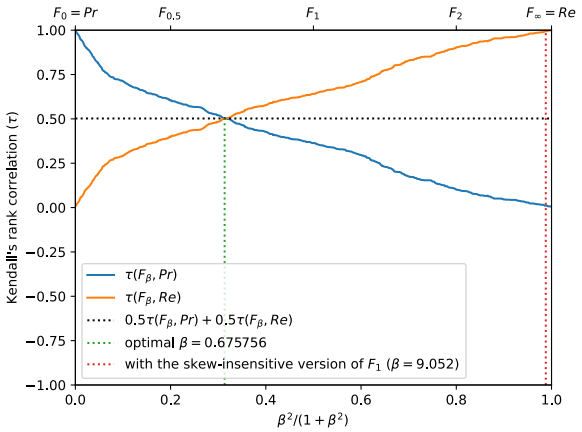


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

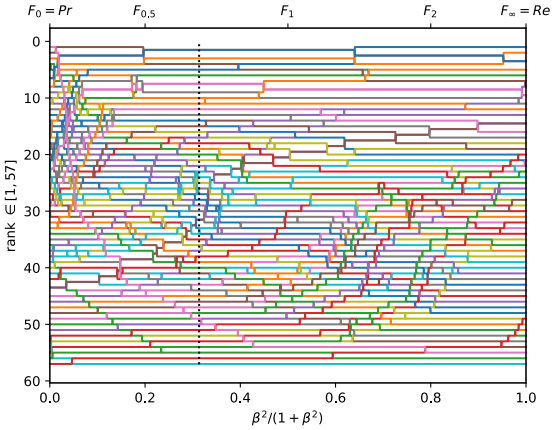
Figure A.3.59. Ranking of 57 BGS methods evaluated on the video "turbulence2" ($\pi_+ = 0.0008$).



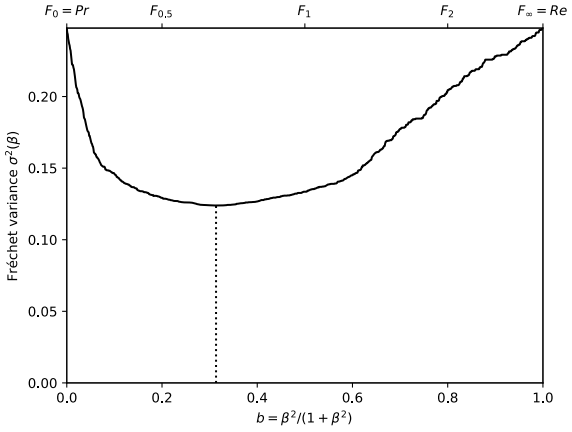
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



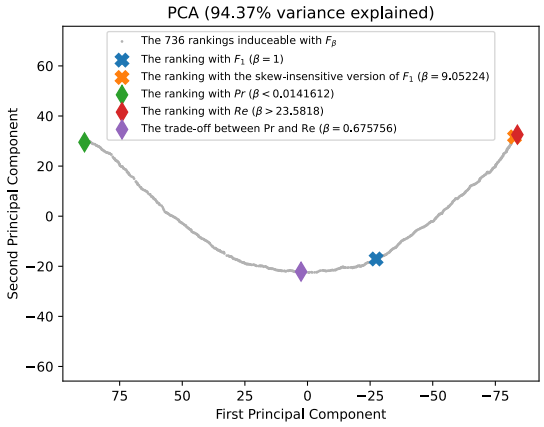
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



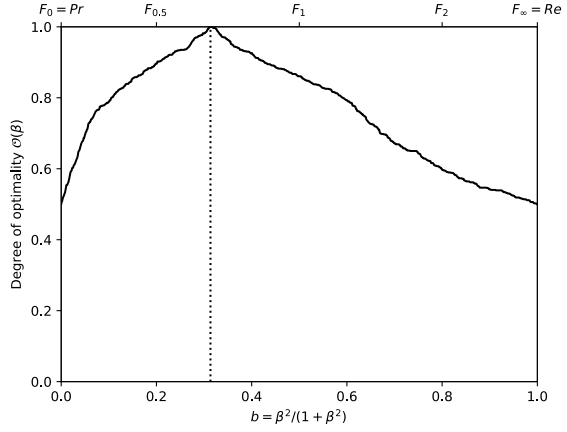
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

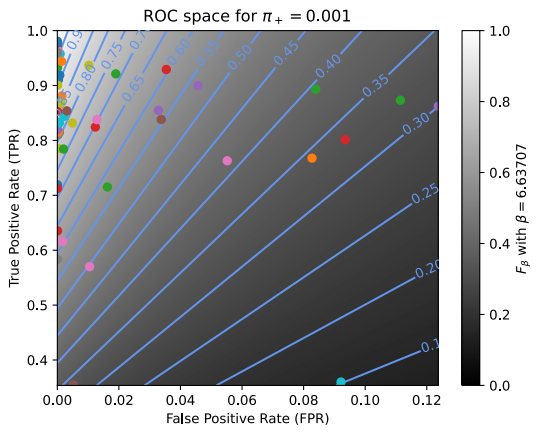


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

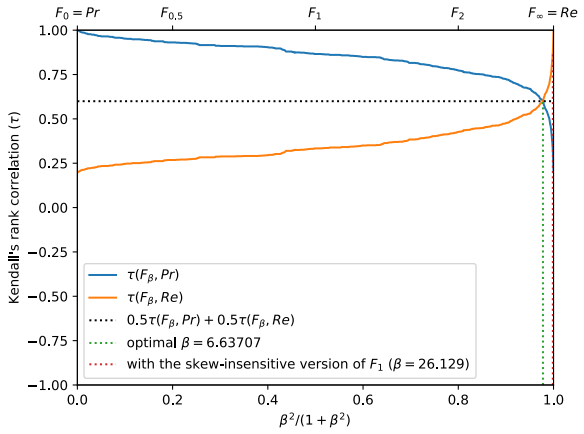


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

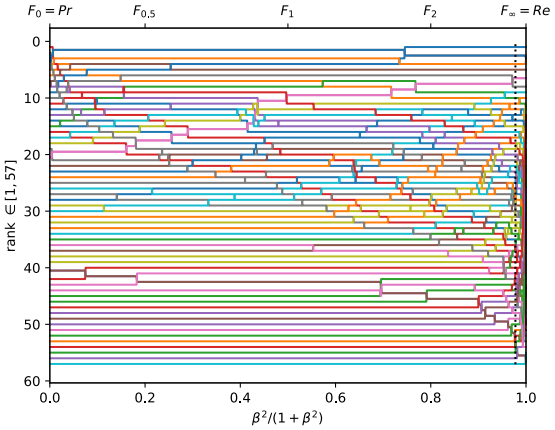
Figure A.3.60. Ranking of 57 BGS methods evaluated on the video "turbulence3" ($\pi_+ = 0.0121$).



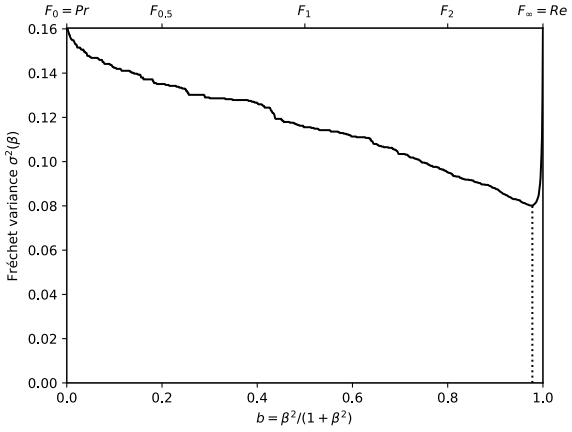
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



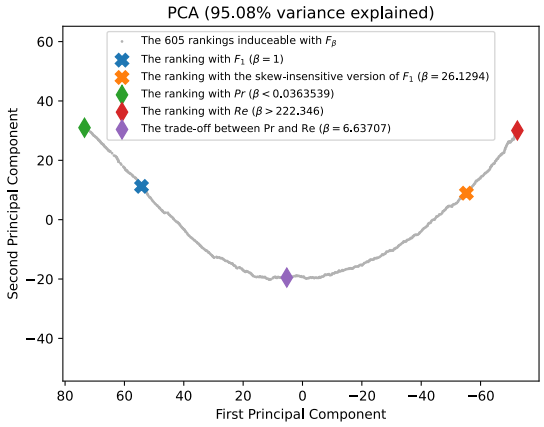
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



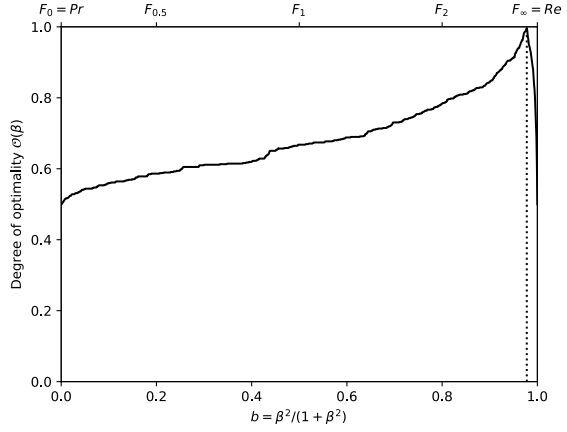
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.

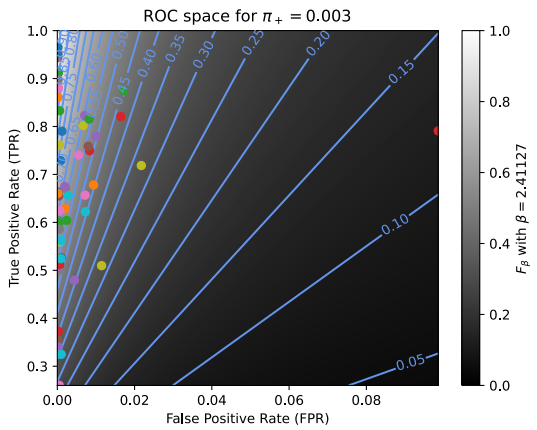


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.

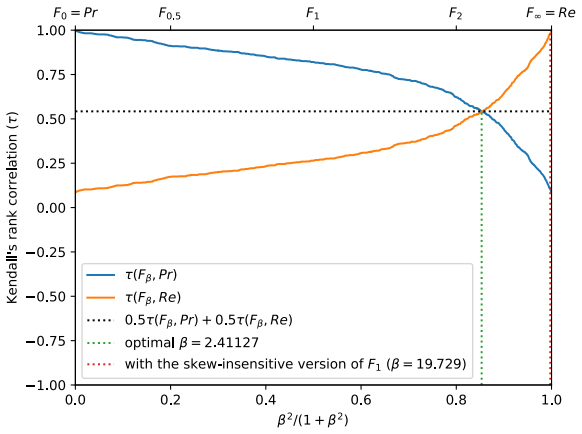


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

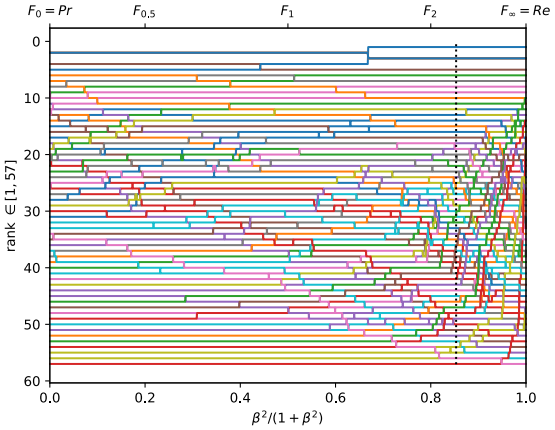
Figure A.3.61. Ranking of 57 BGS methods evaluated on the video "turbulence0" ($\pi_+ = 0.0015$).



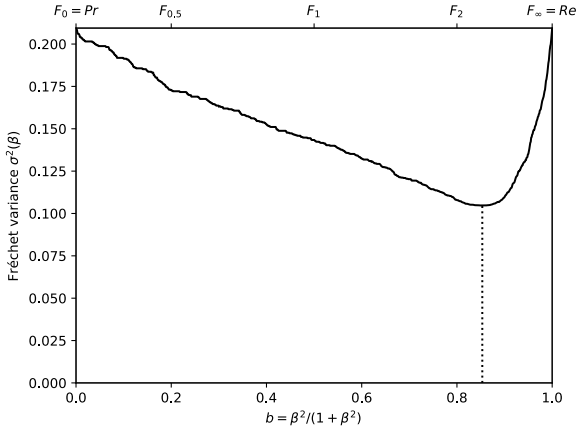
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (12).



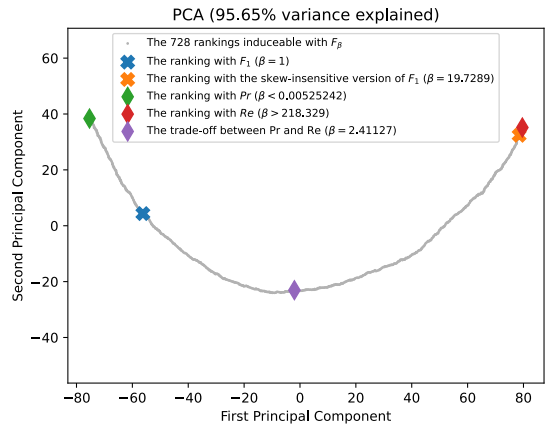
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the two curves intersect.



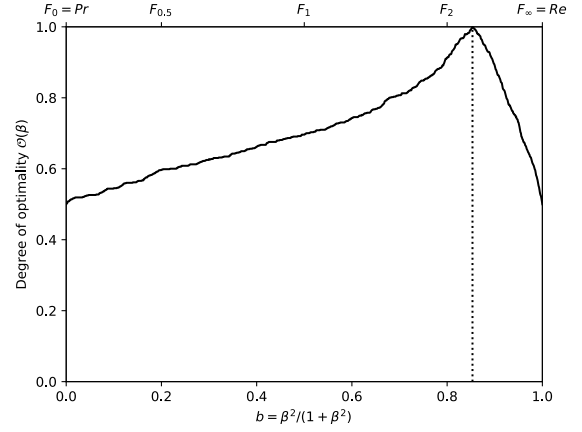
(c) The ranks of each classifier w.r.t. β. The optimal value (or range of optimal values) for β, shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β. The optimal value (or range of optimal values) for β is where the curve has its minimum.



(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1, $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold, with Kendall's distance d_τ.



(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β. It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

Figure A.3.62. Ranking of 57 BGS methods evaluated on the video "turbulence1" ($\pi_+ = 0.0026$).

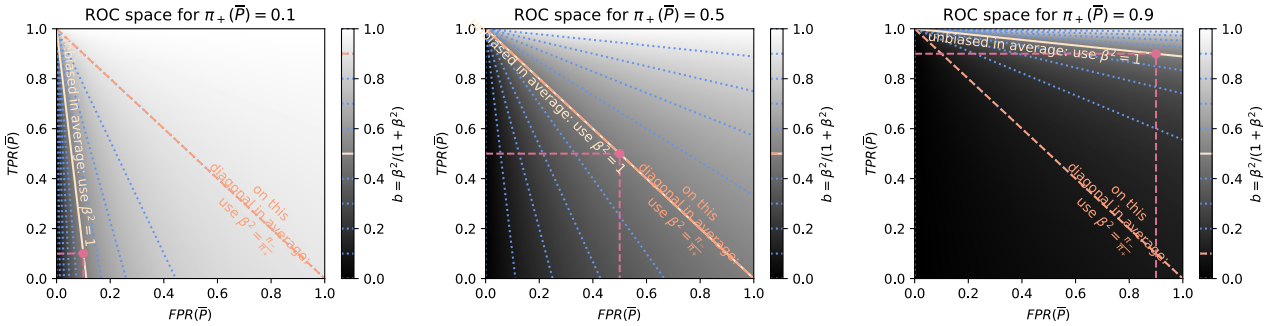


Figure A.3.63. Visualization of our heuristic in ROC space. The recommended value for $b = \frac{\beta^2}{1+\beta^2}$ is the value at the point $(\mathbf{E}[FPR], \mathbf{E}[TPR])$ when the priors are fixed, and the value at the point $(FPR(\overline{P}), TPR(\overline{P}))$ in the general case. This visualization is specific to certain priors: the common priors when the priors are fixed, and $(\pi_-(\overline{P}), \pi_+(\overline{P}))$ in the general case. The more the prior π_+ of the positive class increases, the more the recommended value for β decreases (compare with Fig. 2).

A.3.8. More information on our simple heuristic for ROC users

We provide here additional information about the simple heuristic introduced in Sec. 4.6.

$$\beta^2 = \frac{\mathbf{E}[PFP]}{\mathbf{E}[PFN]} \Leftrightarrow b = \frac{\mathbf{E}[PFP]}{\mathbf{E}[PFN] + \mathbf{E}[PFP]} \tag{70}$$

This heuristic was designed for the typical performances that researchers in the computer vision community (and more specifically those working on "change detection" and "background subtraction") have found useful to report in public benchmarks. For this community, we are fortunate to have a large number of domains (videos) in which a wide range of methods have been evaluated and their performances reported in public rankings. There is however no guarantee that this heuristic performs well in all cases. Nevertheless, as shown in Tab. 1, our heuristic performs reasonably well for all the parametric distributions studied in this paper. It is therefore worth describing it in depth.

As shown in Fig. A.3.63, it is possible to depict the heuristic in the Receiver Operating Characteristic (ROC) space, which is by definition $ROC = (FPR, TPR)$. We start by discussing the particular case in which the class priors are fixed, and then discuss the general case.

- When the class priors (π_-, π_+) are fixed, as $PFP = \pi_- \, FPR$ and $PFN = \pi_+ \, FNR$, our heuristic can be rewritten as

$$\beta^2 = \frac{\pi_- \, \mathbf{E}[FPR]}{\pi_+ \, \mathbf{E}[FNR]} \Leftrightarrow b = \frac{\pi_- \, \mathbf{E}[FPR]}{\pi_+ \, \mathbf{E}[FNR] + \pi_- \, \mathbf{E}[FPR]} \, . \tag{71}$$

  Moreover, as $FNR = 1 - TPR$, we have

$$\beta^2 = \frac{\pi_- \, \mathbf{E}[FPR]}{\pi_+ (1 - \mathbf{E}[TPR])} \Leftrightarrow b = \frac{\pi_- \, \mathbf{E}[FPR]}{\pi_+ (1 - \mathbf{E}[TPR]) + \pi_- \, \mathbf{E}[FPR]} \, . \tag{72}$$

  This last form allows one to see the recommended b (or β) as a function of $\mathbf{E}[FPR]$ and $\mathbf{E}[TPR]$, or in other words as a function of the position of the centroid in ROC.
- In the general case, it is also possible to use the same visualization of the heuristic in ROC. However, instead of looking at the centroid of the performance points, one should look at where the *summarized performance* $\overline{P}$ is projected. Let us remind that, when working with a finite set $\Pi \subset \mathbb{P}_{(\Omega,\Sigma)}$ of performances, the works [28, 30] proposed a consistent way of averaging the score values for all scores. The summarized value for a score X is $X(\overline{P})$, where $\overline{P} \in \mathbb{P}_{(\Omega,\Sigma)}$ is the performance given by $\frac{1}{|\Pi|} \sum_{P \in \Pi} P$. In other words, instead of arithmetically averaging score values, it has been proposed to average, or summarize, performances.
  - For all unconditional probabilistic scores X (such as PFP and PFN), we have $\mathbf{E}[X] = \frac{1}{|\Pi|} \sum_{P \in \Pi} X(P) = X(\overline{P})$. The recommended β depends only on the average, summarized, performance $\overline{P}$.

$$\beta^2 = \frac{PFP(\overline{P})}{PFN(\overline{P})} \Leftrightarrow b = \frac{PFP(\overline{P})}{PFN(\overline{P}) + PFP(\overline{P})} \tag{73}$$

– In general, for the conditional probabilistic scores X, one does not have $X(\overline{P}) = \frac{1}{|\Pi|}\sum_{P\in\Pi} X(P)$. So, $X(\overline{P})$ is not necessarily equal to $\mathbf{E}[X]$. This is the case for TNR, FPR, FNR, and TPR. But as $PFP(\overline{P}) = \pi_-(\overline{P})\,FPR(\overline{P})$ and $PFN(\overline{P}) = \pi_+(\overline{P})\,FNR(\overline{P})$,

$$\beta^2 = \frac{\pi_-(\overline{P})\,FPR(\overline{P})}{\pi_+(\overline{P})\,FNR(\overline{P})} \Leftrightarrow b = \frac{\pi_-(\overline{P})\,FPR(\overline{P})}{\pi_+(\overline{P})\,FNR(\overline{P}) + \pi_-(\overline{P})\,FPR(\overline{P})}\,. \tag{74}$$

Moreover, as $FNR(\overline{P}) = 1 - TPR(\overline{P})$, we have

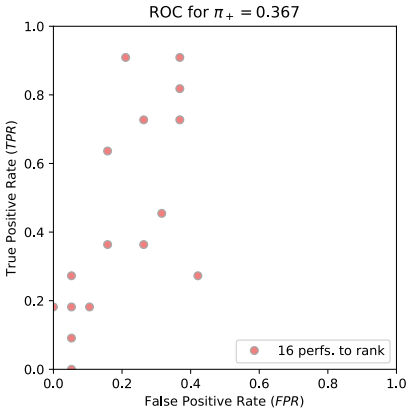
$$\beta^2 = \frac{\pi_-(\overline{P})\,FPR(\overline{P})}{\pi_+(\overline{P})(1 - TPR(\overline{P}))} \Leftrightarrow b = \frac{\pi_-(\overline{P})\,FPR(\overline{P})}{\pi_+(\overline{P})(1 - TPR(\overline{P})) + \pi_-(\overline{P})\,FPR(\overline{P})}\,. \tag{75}$$

It should be stressed that, when the priors differ from one performance to another, the point $(FPR(\overline{P}), TPR(\overline{P}))$ does not coincide with $(\mathbf{E}[FPR], \mathbf{E}[TPR])$. By comparing Eq. (72) with Eq. (75), we see that the same visualization can be used. However, instead of looking at $(\mathbf{E}[FPR], \mathbf{E}[TPR])$, one has to look at $(FPR(\overline{P}), TPR(\overline{P}))$.

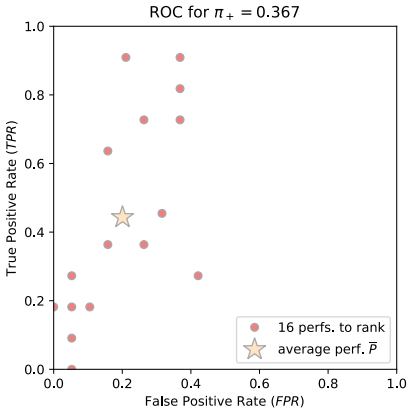
Our heuristic recommends to rank w.r.t. F_1 when the average performance is unbiased (*i.e.*, $Pr = Re$). In the sense of Byrt *et al*. [3], a classifier is unbiased when the classifier predicts as much positive samples as there are in reality. The probability for the predicted class to be positive $P(\{fp, tp\})$ equals the probability for the ground-truth class to be positive $P(\{fn, tp\})$. Thus, $PFP = PFN$ and $Pr = Re$. In ROC, this corresponds to a line passing through $(FPR, TPR) = (0, 1)$ and $(FPR, TPR) = (\pi_+, \pi_+)$. When the performance is unbiased in average, the heuristic recommends $\beta^2 = 1$. This can be seen by taking $\mathbf{E}[PFP] = \mathbf{E}[PFN]$ in Eq. (70) or $PFP(\overline{P}) = PFN(\overline{P})$ in Eq. (73).

Our heuristic recommends to rank w.r.t. $SIVF$ when the priors are fixed, and w.r.t. F_β with $\beta^2 = \pi_-/\pi_+$ in the general case, when the average performance is on the descending diagonal of ROC. If the average performance is on the descending diagonal of ROC, we have $\mathbf{E}[FPR] = 1 - \mathbf{E}[TPR]$ when the priors are fixed, or $FPR(\overline{P}) = 1 - TPR(\overline{P})$ in the general case. Thus, by looking at Eqs. (72) and (75), we see that the recommended value for β^2 is π_-/π_+. When the priors are fixed, the F-score corresponding to this value induces the same performance ordering as the skew-insensitive version of F_1, $SIVF$, defined in [10] as $SIVF = \frac{2TPR}{TPR+FPR+1}$.

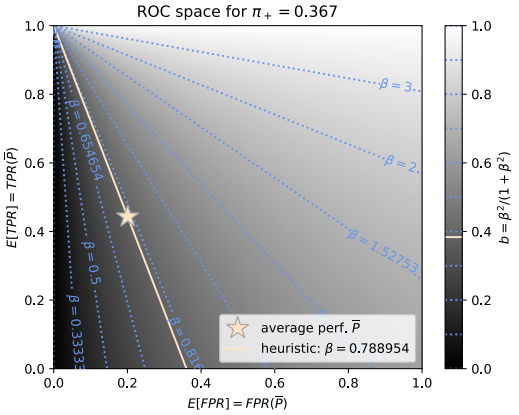
A detailed use of our heuristic on the CADA-RRE example is shown, step by step, in Fig. A.3.64.



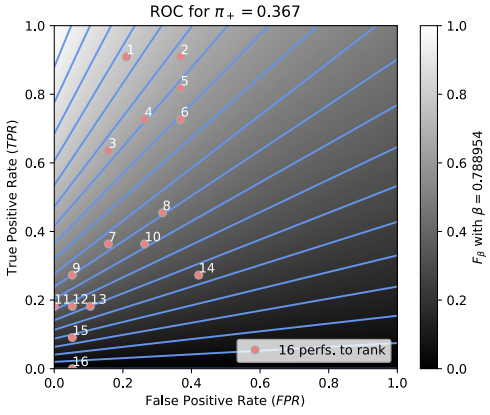
(a) Step 1. Point cloud in the ROC space showing the performances of the methods to rank. Note that all these performances correspond to fixed class priors.



(b) Step 2. Computation of the summarized performance $\overline{P}$ [28]. As all performances correspond to fixed class priors, it is located at the centroid of the point cloud (bisque star).



(c) Step 3. Use of our heuristic. One reads the recommended value for $b = \frac{\beta^2}{1+\beta^2}$ at $(FPR(\overline{P}), TPR(\overline{P}))$.



(d) Step 4. Drawing the score F_β for the recommended β in the background of ROC, with superimposed isometrics depicting the performance ordering induced by it.

Figure A.3.64. Example of use of our heuristic our CADA-RRE example (see Appendix A.2.4).