

What Is the Optimal Ranking Score Between Precision and Recall? We Can Always Find It and It Is Rarely F_1

Sébastien Piérard, Adrien Deliège, and Marc Van Droogenbroeck
 Montefiore Institute, University of Liège, Liège, Belgium
 {S.Pierard,Adrien.Deliège,M.VanDroogenbroeck}@uliege.be

Abstract

Ranking methods or models based on their performance is of prime importance but is tricky because performance is fundamentally multidimensional. In the case of classification, precision and recall are scores with probabilistic interpretations that are both important to consider and complementary. The rankings induced by these two scores are often in partial contradiction. In practice, therefore, it is extremely useful to establish a compromise between the two views to obtain a single, global ranking. Over the last fifty years or so, it has been proposed to take a weighted harmonic mean, known as the F-score, F-measure, or F_β . Generally speaking, by averaging basic scores, we obtain a score that is intermediate in terms of values. However, there is no guarantee that these scores lead to meaningful rankings and no guarantee that the rankings are good tradeoffs between these base scores. Given the ubiquity of F_β scores in the literature, some clarification is in order. Concretely: (1) We establish that F_β -induced rankings are meaningful and define a shortest path between precision- and recall-induced rankings. (2) We frame the problem of finding a tradeoff between two scores as an optimization problem expressed with Kendall rank correlations. We show that F_1 and its skew-insensitive version are far from being optimal in that regard. (3) We provide theoretical tools and a closed-form expression to find the optimal value for β for any distribution or set of performances, and we illustrate their use on six case studies.

1. Introduction

The *precision* Pr (also called *positive predictive value*) and *recall* Re (also called *true positive rate*) are of first importance in classification and other related problems. These scores have a straightforward interpretation, as they give the probabilities of taking the correct decision when the predicted or ground-truth class is positive, respectively.

Unfortunately, working with several scores is often im-

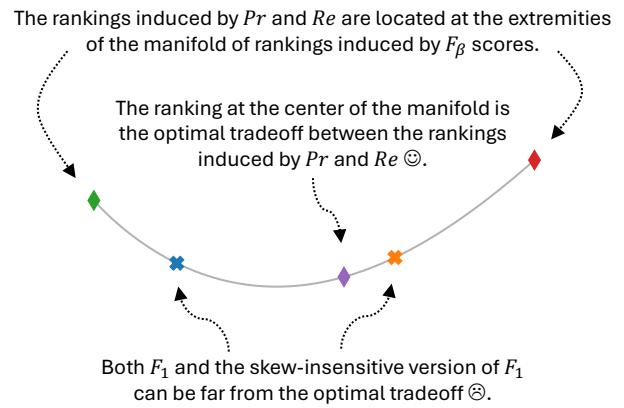


Figure 1. The manifold of rankings inducible with the F_β scores is a curve (drawn as \curvearrowright) that depends on the set or distribution of performances that is considered. The distance along the curve is proportional to the number of swaps of neighboring classifiers needed to transform a ranking into another. The rankings induced by precision \blacklozenge and recall \blacklozenge are at the extremities of the manifold. The rankings induced by the traditional (balanced) $F_1 \blacktriangleleft$ and by the skew-insensitive version of $F_1 \blacktriangleright$ [7] can be anywhere, depending on the set or distribution of performances that are compared, and cannot be considered as the optimal tradeoff between precision and recall. We consider that the ranking located at equal distance, along the curve, from precision and recall is the optimal tradeoff \blacklozenge to rank classifiers.

practical. For this reason, it is common to average the values of Pr and Re . Most often, this is done with a harmonic mean, leading to the family of F-scores F_β , with $\beta \geq 0$:

$$F_\beta^{-1} = (1 - b) Pr^{-1} + b Re^{-1} \quad \text{with } b = \frac{\beta^2}{1 + \beta^2}. \quad (1)$$

These last decades, this family of scores has become very popular in the literature¹. It forms a continuum between pre-

¹With the requests ‘‘F-measure + classification’’, ‘‘F-score + classification’’, and ‘‘F-beta + classification’’, Google Scholar reported on July 23th 2025 that there was about 315,000, 205,000, and 9,080 matching documents, respectively. The score F_1 is also popular in the CVPR community:

cision $Pr = F_0$ and recall $Re = F_\infty$. The traditional (balanced) F-score, F_1 , gives equal weights to Pr and Re . But is F_1 really the optimal tradeoff between precision and recall? The answer to this question depends on the point of view.

The point of view of values. The F-scores are obviously between Pr and Re in the sense that, for any given performance P (*i.e.*, a confusion matrix), the value $F_\beta(P)$ varies monotonically with β , between $Pr(P)$ and $Re(P)$.

The point of view of ranks. This is the point of view that we develop in this paper. Although many authors use the F_1 or F_β scores to rank and compare methods, the claim that they are an appropriate compromise for ranking deserves to be studied. Indeed, several questions arise. How are the rankings induced by the F_β scores spatially organized? Are the F_β scores suitable to induce performance-based rankings? Do the F_β scores induce rankings that form a shortest path between those induced by Pr and Re ? For which $\beta(s)$ do we obtain the optimal tradeoff between the rankings induced by Pr and Re ?

The objective of this paper is to provide answers, teased in Fig. 1, to all these questions. For that purpose, we develop the theory, and we provide experimental results for various distributions and sets of two-class classification performances. We summarize our contributions as follows:

1. We establish that the F_β -family is the right set of scores to consider when looking for an optimal tradeoff between Pr and Re for ranking, as, on the one hand, all F_β lead to meaningful rankings according to the theory of performance-based ranking [20] and, on the other hand, the rankings induced by F_β always form a shortest path between the rankings induced by Pr and Re .
2. We show that neither F_1 , traditionally considered as a balanced compromise between Pr and Re , nor its skew-insensitive version (denoted by $SIVF$ in this paper) introduced in [7] are acceptable tradeoffs for ranking, in general. These scores can lead to rankings far from an optimal one.
3. We provide the theory and methods for defining and finding (with a closed-form expression) the optimal trade-offs between the rankings induced by Pr and Re . These methods are then successfully applied to numerous sets and distributions of performances.

2. Related Work

2.1. Useful References

The meaning of F-scores is discussed in countless papers; see, for example, the work by Christen et al. [4] for a recent review about the F-scores. A plethora of other scores have

we manually estimated that about 10% of the papers accepted at CVPR 2025 refer to or use it.

been defined for classification. They can be found in several reviews [1–3, 5, 6, 12, 13, 21, 22], to cite only a few.

Regarding the ranking of methods or models based on their performances, theoretical foundations have been provided in [20] and the theory has been particularized to the case of two-class crisp classification in [11, 19]. When comparing performances with the same class priors, the isometrics depicted in the ROC (Receiver Operating Characteristic) space is a graphical representation of the performance ordering induced by a given score. The isometrics of Pr and F_1 have been depicted in [7] for different class priors.

There are only a few studies of the correlations between the rankings induced by scores in classification. Both Ferri et al. [6] and Liu et al. [18] used Pearson linear correlations and Spearman rank correlations. To the best of our knowledge, there is no study comparing classification scores based on Kendall rank correlations.

2.2. Building Upon the Related Work

2.2.1. Sometimes, F_1 Mimics the Rankings Induced by Pr or Re and Ignores the Other One

Let us consider the particular case in which the class priors are fixed and denote them by π_- and π_+ for the negative and positive classes, respectively. Extending the work of [7], we depict, in Fig. 2, the isometrics of Pr , F_β for different values of β , Re , and $SIVF$ in the ROC space. For all these scores, we found that the isometrics form pencils of lines with vertices located at $(FPR, TPR) = (-\ell, 0)$ with some $\ell \geq 0$. The performance orderings, and the resulting rankings, depend only on the location of the vertex, and thus on ℓ . For Pr , $\ell = 0$. For Re , $\ell = \infty$. For $SIVF$, $\ell = 1$. And for F_β ,

$$\ell = \beta^2 \frac{\pi_+}{\pi_-} = \beta^2 \frac{\pi_+}{1 - \pi_+}. \quad (2)$$

When $\pi_+ \rightarrow 0$, we see that $\ell \rightarrow 0$ and that all F_β mimic the ranking of $Pr = F_0$ and ignore Re . Similarly, when $\pi_+ \rightarrow 1$, we see that $\ell \rightarrow \infty$ and that all F_β mimic the ranking of $Re = F_\infty$, ignoring Pr . We conclude that *a single value of β cannot provide the optimal ranking for all priors*.

2.2.2. All F_β Scores Lead to Meaningful Rankings, Without Any Constraint on the Performances

There are some conditions that a performance ordering must fulfill to be meaningful. They have been given in the form of axioms in [20]. We found that all these axioms are satisfied by the performance orderings induced by F_β , $\forall \beta \geq 0$. In other words, *all F_β lead to meaningful performance orderings*. It was not a foregone conclusion, because the scores that are, from the point of view of values, between two scores suitable for ranking are not guaranteed to be themselves suitable for ranking. For example, the performance orderings induced by the arithmetic and geometric means of

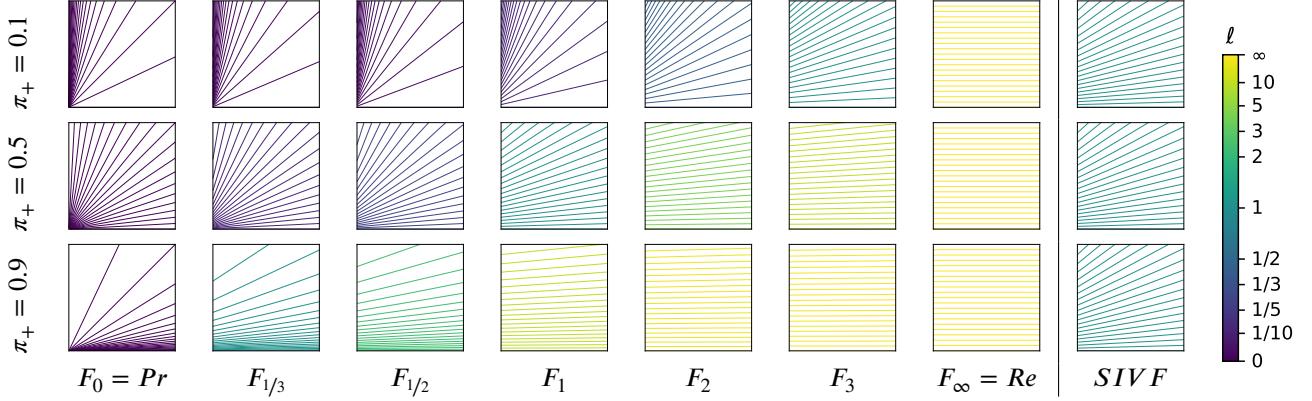


Figure 2. Isometries of F_β and $SIVF$ in the ROC space $(FPR, TPR) \in [0, 1]^2$, for three class priors. An isometric is a locus of equivalent performances, according to the score value. We plotted here those corresponding to the values $0.05, 0.10, 0.15, \dots, 0.95$. The isometries of F_β form a pencil of lines intersecting at $(FPR, TPR) = (-\ell, 0)$ with $\ell = \beta^{2\pi_+/\pi_-}$. The isometries of $SIVF$ intersect at $(FPR, TPR) = (-1, 0)$, no matter what the class priors are. The score $SIVF$ induces the same performance ordering as F_β with $\beta^2 = \pi_-/\pi_+$.

Pr and Re do not satisfy the axioms of performance-based ranking. As the ranking induced by $SIVF$ is the same as the ranking induced by F_β when β^2 is equal to the skew ratio [7] π_-/π_+ , $SIVF$ leads to a meaningful performance ordering when the priors are fixed. However, we found that it does not satisfy the axioms of [20] without this constraint.

3. Theory

3.1. Preliminaries

We use the notations of [20]. $\mathbb{P}_{(\Omega, \Sigma)}$ is the set of all possible performances P in two-class crisp classification, *i.e.* the set of all probability measures with a sample space Ω containing the four cases that can arise: tn (true negative), fp (false positive), fn (false negative), and tp (true positive). The negative and positive class priors are $\pi_- = P(\{tn, fp\})$ and $\pi_+ = P(\{fn, tp\})$, respectively.

We are interested in the rankings induced by scores. Scores are functions $X : \mathbb{P}_{(\Omega, \Sigma)} \rightarrow \mathbb{R} : P \mapsto X(P)$. The unconditional probabilities of a false positive, false negative, and true positive are given by the scores PFN , PFP , and PTP , respectively. The precision (Pr), also called positive predictive value (PPV), is defined as

$$Pr = PPV = F_0 : P \mapsto P(\{tp\} | \{fp, tp\}). \quad (3)$$

The recall (Re), also called true positive rate (TPR), is

$$Re = TPR = F_\infty : P \mapsto P(\{tp\} | \{fn, tp\}). \quad (4)$$

The skew-insensitive version of F_1 [7] is

$$SIVF = \frac{2TPR}{TPR + FPR + 1}, \quad (5)$$

where FPR is the false positive rate.

3.2. The Rankings and Their Manifold

Without loss of generality, let us first consider a finite set $\Pi = \{P_1, \dots, P_n\} \subset \mathbb{P}_{(\Omega, \Sigma)}$ of performances (for infinitely many performances, we just need to consider a distribution² \mathcal{P} instead of a set Π). The rankings of Π are its permutations, in finite number.

Let us consider a score X that induces meaningful rankings. Assuming that the value of X is defined for all $P \in \Pi$, and that X does not assign the same value to two performances of Π (absence of ties), X induces the ranking:

$$\mathbf{x} = (\text{rank}_X(P_1), \dots, \text{rank}_X(P_n)), \quad (6)$$

where $\text{rank}_X(P)$ denotes the number of performances in Π for which X takes a value greater or equal to $X(P)$.

A set of scores (*e.g.* all F_β) corresponds to a set of points as written in Eq. (6). The Euclidean distance (*i.e.* along a straight line) between two rankings is their Spearman distance d_ρ , related to Spearman's ρ [23]. *The distance measured along the manifold of rankings is Kendall's distance d_τ (a.k.a. bubble-sort distance), related to Kendall's τ [16].* We can explain it as follows. Let us create the path graph that connects the rankings that are neighbors. Two rankings \mathbf{x}_a and \mathbf{x}_b are neighbors if and only if they differ only by one swap of two classifiers at consecutive ranks. The geodesic distance, *i.e.* the length of the shortest path, between \mathbf{x}_a and \mathbf{x}_b , is the minimum number of swaps needed to transform \mathbf{x}_a into \mathbf{x}_b , and it is proportional to the Kendall distance d_τ . For this reason, we consider this distance to be adequate for specifying the optimal tradeoff between two rankings.

²Technically, we need to consider a measurable space whose sample space is $\mathbb{P}_{(\Omega, \Sigma)}$. A performance distribution \mathcal{P} then corresponds to a probability measure. Scores, defined as real functions on $\mathbb{P}_{(\Omega, \Sigma)}$, become random variables, and we can rigorously speak of correlation between scores.

As we observed in all our case studies that the manifold is curved, as depicted in Fig. 1, we will define the optimal tradeoff in terms of Kendall's distance d_τ or in terms of Kendall's rank correlation τ .

3.3. The Shortest Paths and the Means of Scores

The score F_β is the weighted harmonic mean between precision Pr and recall Re . The harmonic mean is the generalized f -mean, also called the regular mean [17], with $f : x \mapsto x^{-1}$. This has an important implication for the performance orderings induced by Pr , Re , and F_β .

Consider an interval $\xi \subseteq \mathbb{R}$, m scores X_1, \dots, X_m whose images are included in ξ , and a continuous strictly monotonic real function f defined on ξ . If a performance P_A is better than, equivalent to, or worse than a performance P_B according to all the scores X_1, \dots, X_m , then so is it according to all their generalized f -means \bar{X} . This extends to weighted f -means. So, *if a performance P_A is better than, equivalent to, or worse than a performance P_B according to Pr and to Re , then so is it according to all the F_β scores.*

Let us introduce the indicator $\Delta_{X_1, X_2}^{P_A, P_B} \in \{0, 1\}$ that specifies if the scores X_1 and X_2 disagree on the relative order between the performances P_A and P_B . Given the implication that a mean of scores has on the rankings, we have

$$\Delta_{Pr, Re}^{P_A, P_B} = \Delta_{Pr, F_\beta}^{P_A, P_B} + \Delta_{F_\beta, Re}^{P_A, P_B} \quad \forall \beta \geq 0. \quad (7)$$

And, as [16]

$$d_\tau(X_1; X_2) = \frac{2}{n(n-1)} \sum_{i < j} \Delta_{X_1, X_2}^{P_i, P_j} \in [0, 1], \quad (8)$$

$$d_\tau(Pr; Re) = d_\tau(Pr; F_\beta) + d_\tau(F_\beta; Re) \quad \forall \beta \geq 0. \quad (9)$$

This means that *the rankings induced by the F_β scores form one shortest path (i.e. a geodesic) between the rankings induced by Pr and Re* . This is a key result for this paper, as it proves meaningful the search for the optimal tradeoff between Pr and Re in the family of F_β scores. As $\tau = 1 - 2d_\tau$, Eq. (9) can be rewritten in terms of correlations as

$$1 + \tau(Pr; Re) = \tau(Pr; F_\beta) + \tau(F_\beta; Re) \quad \forall \beta \geq 0. \quad (10)$$

3.4. The Optimal Tradeoff Between Pr and Re

3.4.1. Optimal Tradeoffs as Karcher Means

The optimal tradeoffs are the F_β scores that minimize the Fréchet variance [8]:

$$\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re). \quad (11)$$

The solutions $F_\beta = F_*$ of Eq. (11), known as the Karcher means [15] are those that are equidistant of Pr and Re (this

is detailed in supplementary material), *i.e.* such that

$$d_\tau(Pr; F_*) = d_\tau(F_*; Re) = \frac{d_\tau(Pr; Re)}{2} \quad (12)$$

$$\Leftrightarrow \tau(Pr; F_*) = \tau(F_*; Re) = \frac{1 + \tau(Pr; Re)}{2}. \quad (13)$$

3.4.2. A Closed-Form Expression for the Optimal β

Let us, by thought, place ourselves at the ranking induced by $Pr = F_0$ and then continuously increase β , following the path of rankings induced by the F_β scores, towards the one induced by $Re = F_\infty$. With a finite number of performances, we stay at a given ranking during a range of values for β and then, suddenly, move to the next ranking. The value of β at which such a transition occurs is a value for which two performances are put on an equal footing by F_β . Let us consider any two performances P_1 and P_2 . We have $F_\beta(P_1) = F_\beta(P_2)$ if and only if $\beta^2 = \vartheta(P_1, P_2)$ with

$$\vartheta(P_1, P_2) = -\frac{PTP(P_1)PFP(P_2) - PTP(P_2)PFP(P_1)}{PTP(P_1)PFN(P_2) - PTP(P_2)PFN(P_1)}, \quad (14)$$

when this value is positive. Otherwise, there is no score F_β that puts the performances P_1 and P_2 in equivalence. Kendall's distance between Pr and F_β increases linearly with the number of swaps. This leads us to the conclusion that *the optimal tradeoff is $F_* = F_\beta$ with*

$$\beta^2 = \text{median} (\{\vartheta(P_i, P_j) \mid i \neq j \wedge \vartheta(P_i, P_j) \geq 0\}). \quad (15)$$

3.4.3. Assessing the Degree of Optimality of Some F_β

The choice of β comes down to choosing a permutation of the classifiers. Looking at the level of pairwise comparisons, we can identify three mutually exclusive cases that we denote by \boxtimes , \boxtimes , and \checkmark .

\boxtimes : the pairs of classifiers that are ordered in the same way by Pr and Re . As it is also the case for all β , there is no choice to make. The proportion of such pairs is given by

$$\begin{aligned} P(\boxtimes) &= 1 - d_\tau(Pr; Re) \\ &= 1/2 [1 + \tau(Pr; Re)] = 1/2 [\tau(Pr; F_\beta) + \tau(F_\beta; Re)]. \end{aligned} \quad (16)$$

\boxtimes : the pairs of classifiers for which a choice has to be made because Pr and Re contradict each other, and for which the choice is not optimal (F_β disagrees with F_*). The proportion of such pairs is given by

$$P(\boxtimes) = d_\tau(F_\beta; F_*) = 1/4 |\tau(Pr; F_\beta) - \tau(F_\beta; Re)|. \quad (17)$$

\checkmark : the pairs of classifiers for which a choice has to be made and for which the choice is optimal. The proportion of such pairs is given by

$$P(\checkmark) = d_\tau(Pr; Re) - d_\tau(F_\beta; F_*) = 1 - P(\boxtimes) - P(\boxtimes). \quad (18)$$

It appears from Eqs. (16) and (17) that *if one uses some F_β to rank and reports the values for both $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$, then we can determine the degree of optimality for the chosen β* . We define it as follows:

$$\mathcal{O} = P(\checkmark | \bar{\Xi}) = \frac{P(\checkmark)}{P(\checkmark) + P(\bar{\Xi})} = 1 - \frac{P(\bar{\Xi})}{1 - P(\bar{\Xi})}. \quad (19)$$

Note that $\mathcal{O} = 1$ if and only if $\tau(Pr; F_*) = \tau(F_*; Re)$, which is what is targeted by minimizing the Fréchet variance defined in Eq. (11).

4. Case Studies

We now perform case studies for several distributions (see Fig. 3) and many sets of performances. The results have been obtained either analytically (with the help of Wolfram 14.2 software; all details in supplementary material) or numerically based on Monte Carlo simulations. A summary of the results, emphasizing the degree of optimality and providing the links to results that can be used to determine the optimal tradeoff, is provided in Tab. 1.

4.1. Uniform Distribution Over All Performances

Let us start by considering the uniform distribution over the set Π_1 of all performances:

$$\Pi_1 = \mathbb{P}_{(\Omega, \Sigma)}. \quad (20)$$

It is the Dirichlet distribution with all concentration parameters set to 1, and corresponds to a uniform distribution in a tetrahedron (see Fig. 3a). With this distribution, precision and recall have uniformly distributed values and are positively correlated: $\tau(Pr; Re) = 1/3$.

The F_1 score has not uniformly distributed values, but is the optimal tradeoff between Pr and Re :

$$\tau(Pr; F_1) = \tau(F_1; Re) = 2/3. \quad (21)$$

The skew-insensitive version of F_1 , $SIVF$, does not lead to meaningful rankings as it does not satisfy the 3rd axiom of the theory of performance-based ranking [20] on Π_1 . Moreover, we computed $\tau(Pr; SIVF) \approx 0.43$ and $\tau(SIVF; Re) \approx 0.81$, which implies that $SIVF$ it is not located on a shortest path between Pr and Re and is not at equidistance from them.

4.2. Uniform Distributions With Fixed Probability of True Negatives

Motivated by the fact that Pr , Re , and F_β give no importance to true negatives, we now consider a second type of distributions, the uniform distribution over the sets

$$\Pi_2(ptn) = \{P \in \mathbb{P}_{(\Omega, \Sigma)} : P(\{tn\}) = ptn\} \subset \Pi_1. \quad (22)$$

Note that $\Pi_2(0)$ is the set of performances considered in detection problems. As shown in Fig. 3b, these distributions

correspond to uniform distributions over equilateral triangles. Pr and Re have uniformly distributed values and are positively correlated: $\tau(Pr; Re) = 1/3$.

For F_1 , our observations remain the same as what we had without any constraint, *i.e.* it is the optimal tradeoff:

$$\tau(Pr; F_1) = \tau(F_1; Re) = 2/3. \quad (23)$$

For $SIVF$, we found that it does not satisfy the 3rd axiom of the theory of performance-based ranking [20] on $\Pi_2(ptn)$, unless $ptn = 0$. But, to the contrary of what we had without any constraint, $SIVF$ is now located on a shortest path between Pr and Re . However, it is still far from being at equidistance from them. For $ptn = 0$, we found $\tau(Pr; SIVF) = 1/3$ and $\tau(SIVF; Re) = 1$.

4.3. Uniform Distributions With Fixed Class Priors

We now move to more realistic distributions for classification, and consider that one needs to rank only performances corresponding to some given class priors (π_- , π_+). For this reason, we consider the uniform distributions over the sets

$$\Pi_3(\pi_+) = \{P \in \mathbb{P}_{(\Omega, \Sigma)} : P(\{fn, tp\}) = \pi_+\} \subset \Pi_1. \quad (24)$$

As shown in Fig. 3c, these distributions correspond to uniform distributions over rectangles. It turns out that their axes correspond to $FPR \in [0, 1]$ and to $TPR \in [0, 1]$. The uniform distributions over $\Pi_3(\pi_+)$ correspond thus to uniform distributions of points in ROC. We found, no matter what the class priors are, that $\tau(Pr; Re) = 1/2$.

The manifold can be easily visualized by PCA. For example, if we apply a two-components PCA to the rankings obtained with $\Pi_3(\pi_+)$ for $\pi_+ = 0.1$, we obtain the curve depicted in Fig. 1 (98.77% of total variance explained).

We stress that, to the contrary of what we concluded for the two previous distributions, when the class priors are fixed, F_1 is far from the optimal tradeoff, as shown in Fig. 4a: unless $\pi_+ \approx 0.381$, $\tau(Pr; F_1) \neq \tau(F_1; Re)$. In the extreme cases for which the prior of the negative (positive) class tends towards 1.0, the ranking induced by F_1 perfectly mimics the ranking induced by Pr (Re), thus totally ignoring the ranking induced by Re (Pr).

To find the optimal threshold, we need first to compute either $\tau(Pr; F_\beta)$ or $\tau(F_\beta; Re)$. The other can be found using Eq. (10). Analytically, we found

$$\tau(F_\beta; Re) = \frac{1}{2} + \ell - \ell^2 \log \frac{1 + \ell}{\ell}, \quad (25)$$

with ℓ defined as in Eq. (2). Then, we minimized the Fréchet variance, given in Eq. (11). If one considers some given class priors, then it is a function of β (see Fig. 4b) and the minimization leads to the optimal value of β for $F_* = F_\beta$. However, it is also possible to see the Fréchet variance as a function of ℓ . Its minimization leads to the solution $\ell \approx$

Case study	$\tau(Pr; Re)$	$F_1 (\beta^2 = 1)$	$SIVF (\equiv \beta^2 = \frac{\pi_-}{\pi_+})$	Heuristic $\beta^2 = \frac{E[PFP]}{E[PN]}$	Optimal tradeoff, $\mathcal{O} = 100\%$
uniform distribution over all performances	$1/3$	$\mathcal{O} = 100\% \text{ (optimal)}$	NA (meaningless ranking)	$\mathcal{O} = 100\% \text{ (optimal)}$ (it selects F_1)	our analytical result cf. Eq. (21)
uniform distributions with fixed probability of true negatives	$1/3$	$\mathcal{O} = 100\% \text{ (optimal)}$	NA (meaningless ranking)	$\mathcal{O} = 100\% \text{ (optimal)}$ (it selects F_1)	our analytical result cf. Eq. (23)
uniform distributions with fixed class priors	$1/2$	$\mathcal{O} \in [50\%, 100\%]$ optimal only for $\pi_+ \simeq 0.381$	$\mathcal{O} = \log(4) - 1/2 \simeq 88.63\%$	$\mathcal{O} = \log(4) - 1/2 \simeq 88.63\%$ (it selects $SIVF$)	our analytical result cf. Eq. (26) and Fig. 4c
uniform distributions with fixed class priors, above no-skill	0	$\mathcal{O} \in [50\%, 100\%]$ optimal only for $\pi_+ \simeq 0.325$	$\mathcal{O} = 5/6 \simeq 83.33\%$	$\mathcal{O} = 5/6 \simeq 83.33\%$ (it selects $SIVF$)	our analytical result cf. Eq. (28)
uniform distributions with fixed class priors, close to oracle	$\in (0, 1/2)$	$\mathcal{O} \in [87.85\%, 100\%]$ optimal only for $\pi_+ \rightarrow 1$	$\mathcal{O} \in [50\%, 100\%]$ optimal only for $\pi_+ \simeq 0.561$	$\mathcal{O} \in [87.85\%, 100\%]$ (it selects F_1)	our numerical result cf. Fig. 5c
53 real sets of about 60 performances	$\in [-0.31, 0.65]$	$\mathcal{O} \in [52.23\%, 100.00\%]$ (mean: 78.44%)	$\mathcal{O} \in [50.00\%, 99.13\%]$ (mean: 56.71%)	$\mathcal{O} \in [69.69\%, 100.00\%]$ (mean: 89.53%)	our closed-form expression for the optimal β , Eq. (15)

Table 1. Summary of our results emphasizing the degree of optimality \mathcal{O} introduced in Eq. (19). As shown by the last column, we can always find the optimal ranking score (in green) between Pr and Re . Relying on a fixed and arbitrarily chosen score like F_1 or $SIVF$ can lead to catastrophic cases (in red). In our case studies, the simple heuristic introduced in Sec. 4.6 never led to such catastrophic cases, but nevertheless appeared to be suboptimal in some cases (in orange).

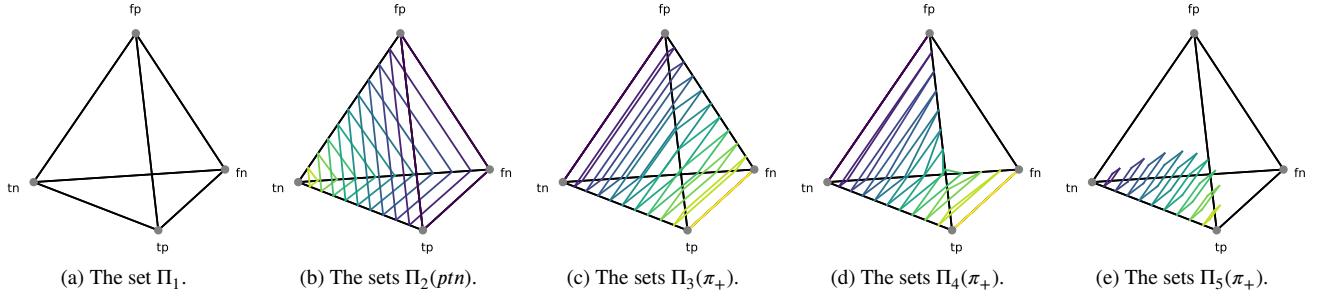


Figure 3. Graphical representation of some sets of two-class crisp classification performances considered in this paper. (a) All possible performances (see Eq. (20)) correspond to points in a regular tetrahedron. (b) The performances corresponding to fixed probabilities of true negatives (see Eq. (22)) are points in parallel equilateral triangles. (c) The performances corresponding to fixed class priors (see Eq. (24)) are points in parallel rectangles. (d) The performances corresponding to fixed class priors and “above” no-skills (see Eq. (27)) are points in parallel right triangles. (e) The performances for given class priors, and “close to the oracle” (see Eq. (29)) are points in parallel squares. The color represents the value of the fixed parameter considered (either ptn or π_+).

0.61585. Injecting this value into Eq. (2), we obtained the link between the optimal β and the class priors:

$$\beta^2 = \ell \frac{\pi_-}{\pi_+} \Rightarrow b \simeq \frac{0.61585 (1 - \pi_+)}{\pi_+ + 0.61585 (1 - \pi_+)} \quad (26)$$

It is depicted in Fig. 4c. We see that the optimal relationship is quite close to that of $SIVF$ (*i.e.*, $b = 1 - \pi_+$), but very different from that of F_1 (*i.e.*, $b = 1/2$).

4.4. Uniform Distributions With Fixed Class Priors, Above No-Skill

Performances below the rising diagonal of ROC, where the no-skill performances lie, are usually of little interest. Systematically inverting the output of a classifier is equivalent to applying a central symmetry in ROC. Thus, from any performance below the no-skills, it is easy to obtain a performance above the no-skills. But the rankings induced by Pr , Re , and F_β are sensitive to this operation. For this reason,

we study the case of the uniform distributions over the sets

$$\begin{aligned} \Pi_4(\pi_+) = \{P \in \mathbb{P}_{(\Omega, \Sigma)} : P(\{fn, tp\}) = \pi_+ \\ \wedge TPR(P) \geq FPR(P)\} \subset \Pi_3(\pi_+). \end{aligned} \quad (27)$$

These sets are depicted in Fig. 3d. No matter what the class priors are, we obtain $\tau(Pr; Re) = 0$. All detailed results are in supplementary material. In a nutshell, the observations are very close to those reported for the uniform distributions with fixed class priors. The overall shape of the manifold, as observed by PCA, is very similar. $\tau(Pr; F_1)$ monotonically increases from 0 to 1 (instead of 0.5 to 1) and $\tau(F_1; Re)$ monotonically decreases from 1 to 0 (instead of 1 to 0.5) when π_+ sweeps the $[0, 1]$ interval. The F_1 score is the optimal tradeoff only for $\pi_+ \simeq 0.325$. Regarding the adaptation, the curve is slightly more curved and lower than the one we depicted in Fig. 4c, with the optimal value $\ell \simeq 0.48$:

$$\beta^2 = \ell \frac{\pi_-}{\pi_+} \Rightarrow b \simeq \frac{0.48 (1 - \pi_+)}{\pi_+ + 0.48 (1 - \pi_+)} . \quad (28)$$

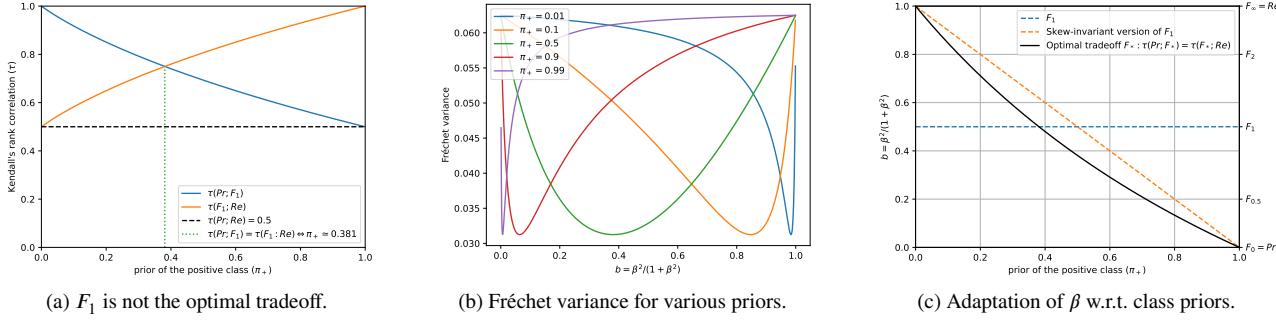


Figure 4. Results for uniform distributions over the performances with fixed class priors, *i.e.* $\Pi_3(\pi_+)$.

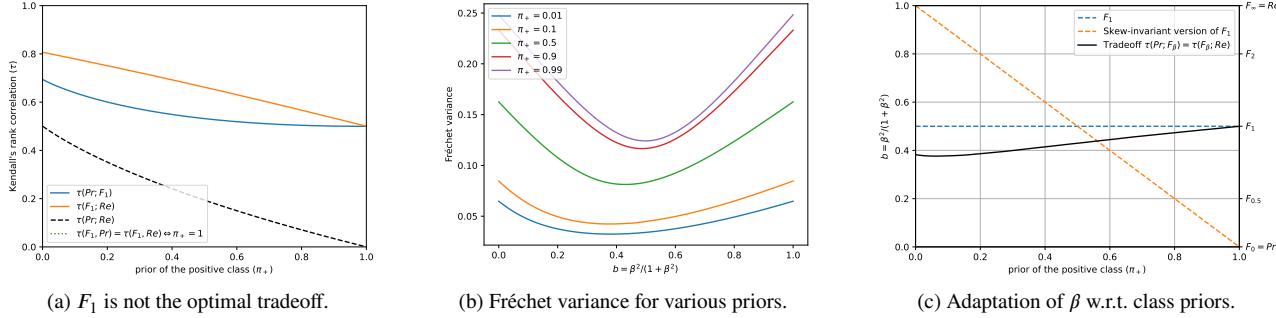


Figure 5. Results for uniform distributions over the performances with fixed class priors and close to the oracle, *i.e.* $\Pi_5(\pi_+)$.

4.5. Uniform Distributions With Fixed Class Priors, Close to Oracle

In competitions, a setting in which ranking is omnipresent, contenders usually refrain from choosing decision thresholds that would lead either to a high TPR at the cost of a high FPR or to a low FPR at the cost of a low TPR . In our opinion, this means that $FPR < \pi_+$ and $TPR > \pi_+$. In ROC, these performances are above and on the left-hand side of the no-skill performance that results from randomly choosing the negative or positive class with respective probabilities equal to π_- and π_+ . For this reason, we now discuss the case of uniform distributions over the sets

$$\begin{aligned} \Pi_5(\pi_+) &= \{P \in \mathbb{P}_{(\Omega, \Sigma)} : P(\{fn, tp\}) = \pi_+ \\ &\wedge FPR(P) < \pi_+ \wedge TPR(P) > \pi_+\} \subset \Pi_4(\pi_+). \end{aligned} \quad (29)$$

As shown in Fig. 3e, these distributions correspond to uniform distributions over squares. Drawing a performance at random can be achieved by drawing independent values uniformly in $[0, \pi_+]$ for FPR and in $[\pi_+, 1]$ for TPR .

The results, shown in Fig. 5, are remarkably different from the previous ones. In comparison to Fig. 4, the optimal β now increases from about 0.8 to 1, when π_+ increases from 0 to 1, while it was previously decreasing from ∞ to 0.

4.6. Some Real Sets of Performances

We now report results for performances encountered in existing rankings. They are either performances of some baseline methods or performances that people considered good enough for a competition. More specifically, we considered a computer vision task known as *background subtraction*, similar to a pixelwise classification between background and foreground [9, 14]. The results of about 60 methods (classifiers), on 53 videos—a dataset known as *CDnet 2014* [24]—are publicly available on changedetection.net. According to [10], the multi-criteria ranking proposed by this platform is well correlated with F_1 . On any given video, all performances belong to $\Pi_3(\pi_+)$. Moreover, 99.9% and 99.4% of them belong to $\Pi_4(\pi_+)$ and $\Pi_5(\pi_+)$, respectively. Depending on the video, $\tau(Pr; Re) \in [-0.31, 0.65]$.

Figure 6 shows the results for the set of performances reported for one video; more results are to be found in the supplementary material. The optimal β is between 0.29 and 22.79 depending on the video. It is in the $[0.5, 2]$ range for only 35 videos among the 53.

An interesting observation, shown in Fig. 7, is that the class priors (π_-, π_+) are poor predictors for the optimal value of β . However, b is well correlated with

$$\frac{\mathbf{E}[PFP]}{\mathbf{E}[PFN] + \mathbf{E}[PFP]} = \frac{\sum_{P \in \Pi} P(\{fp\})}{\sum_{P \in \Pi} P(\{fn, fp\})}, \quad (30)$$

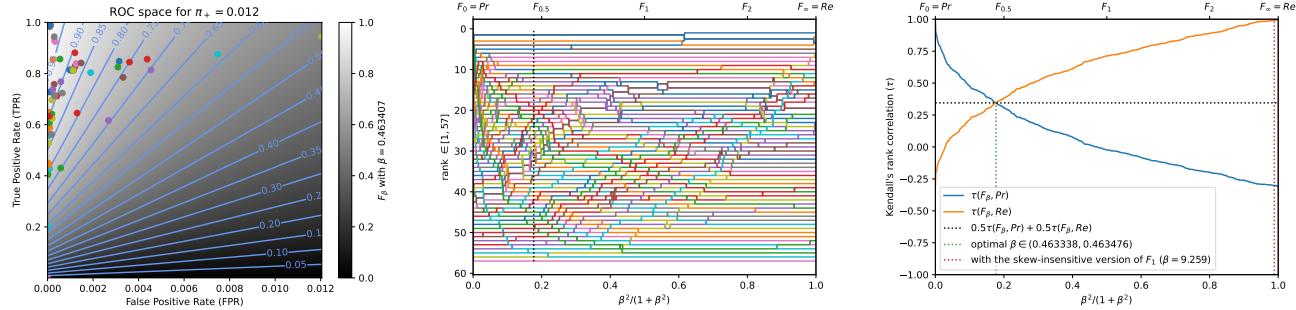


Figure 6. What is the optimal tradeoff for the performances of 57 background subtraction methods (pixelwise classifiers) on the *blizzard* video of *CDnet 2014*? Left: the 57 performances to be ranked, depicted in the ROC space, with the isometrics of the optimal tradeoff. Center: the ranks of each classifier, w.r.t. β . Right: $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The ranking with F_1 , at the center, is too far from Pr and too close to Re . The optimal tradeoff, for which $\tau(Pr; F_\beta) = \tau(F_\beta; Re)$ is at $\beta \simeq 0.463$.

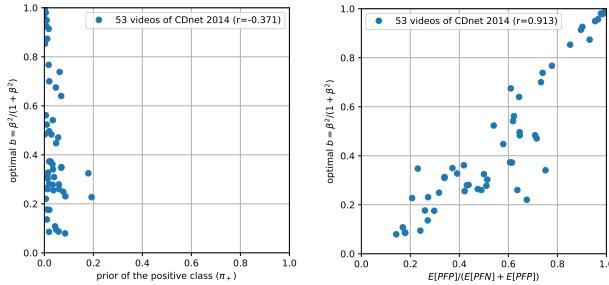


Figure 7. Which statistic about the sets of performances can be used to predict the optimal value of β ? These results are based on 53 sets of performances (those of about 60 methods on each video of *CDnet 2014*). Left: the class priors are not suitable. Right: the ratio of the average probability of false positives ($E[PFP]$) to the average probability of errors ($E[PFN] + E[PFP]$) seems to be a good predictor for β , as Pearson's linear correlation between $\frac{E[PFP]}{E[PFN]+E[PFP]}$ and $\frac{\beta^2}{1+\beta^2}$ is $r \simeq 0.913$.

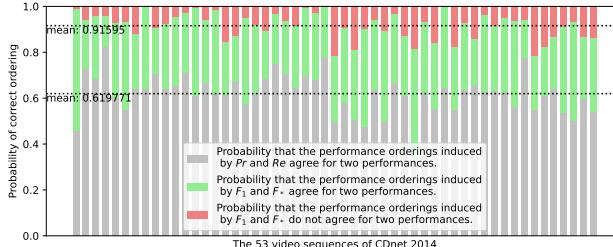


Figure 8. How far is F_1 from the optimal tradeoff F_* ? Each vertical bar is for a specific set of performances. ■: $P(\text{X})$, ■: $P(\text{V})$, ■: $P(\text{X})$.

which suggests that $\beta^2 = \frac{E[PFP]}{E[PFN]}$ is a simple heuristic that is worth trying as an alternative to the closed-form and optimal solution provided in Eq. (15). For the sake of completeness, the degree of optimality \mathcal{O} achieved by this heuristic is provided in Tab. 1 for all our case studies.

Figure 8 shows the consequences of using F_1 for ranking instead of the optimal tradeoff F_* . On average, there is 61.98% (■) of chance that a randomly chosen pair of methods is ranked the same by Pr and Re . In this case, which tradeoff is chosen among all F_β is not relevant. But when Pr and Re disagree, the value of β matters. The degree of optimality for F_1 is only about $\mathcal{O} \simeq 78\%$ (■) on average.

5. Conclusion

Nobody would ever choose a classifier solely based on the precision Pr or the recall Re . The traditional approach of calculating the harmonic mean F_1 of these two scores and then ranking the classifiers based on it is, unfortunately, nothing but a red herring. The colossal problem, ignored in the literature until now, is that the ranking built upon the balanced average of two scores is generally not a good trade-off between the rankings obtained with the two base scores. In extreme cases, the ranking obtained with F_1 copies the ranking based on Pr or Re and totally ignores the other.

To solve these issues, we defined and described the manifold of rankings induced by scores (see Fig. 1) and found that the distance along it is linearly related to Kendall's rank correlation τ . Besides, we showed that all F_β scores induce meaningful rankings, even without any constraint on the performances. We further proved that the F_β scores induce rankings that form a shortest path between the rankings of Pr and Re . Also, we provided the theory and methods (including a closed-form expression) for defining and identifying the optimal tradeoffs between the rankings induced by Pr and Re . Lastly, we exemplified six case studies, covering numerous distributions and sets of performances. A summary of our results is provided in Tab. 1.

Acknowledgments. S. Piérard is funded by grants 8573 (ReconnAissance project) and 2010235 (ARIAC by DIGITALWALLONIA4.AI) of the SPW EER, Wallonia, Belgium; A. Deliège is a F.R.S.-FNRS postdoc researcher.

References

- [1] Davide Ballabio, Francesca Grisoni, and Roberto Todeschini. Multivariate comparison of classification performance measures. *Chemometrics and Intelligent Laboratory Systems*, 174:33–44, 2018. 2
- [2] Gurol Canbek, Seref Sagiroglu, Tugba Taskaya Temizel, and Nazife Baykal. Binary classification performance measures/metrics: A comprehensive visualized roadmap to gain new insights. In *International Conference on Computer Science and Engineering (UBMK)*, pages 821–826, Antalya, Turkey, 2017. Institute of Electrical and Electronics Engineers (IEEE).
- [3] Seung-Seok Choi, Sung-Hyuk Cha, and Charles C. and Tapert. A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics*, 8(1):43–48, 2010. 2
- [4] Peter Christen, David J. Hand, and Nishadi Kirielle. A review of the F-measure: Its history, properties, criticism, and alternatives. *ACM Computing Surveys*, 56(3):1–24, 2023. 2
- [5] Luciana Ferrer. Analysis and comparison of classification metrics. *arXiv*, abs/2209.05355, 2022. 2
- [6] César Ferri, José Hernández-Orallo, and Ramona Modroiu. An experimental comparison of performance measures for classification. *Pattern Recognition Letters*, 30(1):27–38, 2009. 2
- [7] Peter A. Flach. The geometry of ROC space: Understanding machine learning metrics through ROC isometrics. In *International Conference on Machine Learning (ICML)*, pages 194–201, Washington, District of Columbia, USA, 2003. 1, 2, 3, 11, 22, 25, 28
- [8] Maurice Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'institut Henri Poincaré*, 10(4):215–310, 1948. 4, 13
- [9] Belmar Garcia-Garcia, Thierry Bouwmans, and Alberto Jorge Rosales Silva. Background subtraction in real applications: Challenges, current models and future directions. *Computer Science Review*, 35:1–42, 2020. 7
- [10] Nil Goyette, Pierre-Marc Jodoin, Fatih Porikli, Janusz Konrad, and Prakash Ishwar. Chagedetection.net: A new change detection benchmark dataset. In *IEEE International Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1–8, Providence, RI, USA, 2012. Institute of Electrical and Electronics Engineers (IEEE). 7
- [11] Anaïs Halin, Sébastien Piérard, Anthony Cioppa, and Marc Van Droogenbroeck. A hitchhiker’s guide to understanding performances of two-class classifiers. *arXiv*, abs/2412.04377, 2024. 2
- [12] Mohammad Hossin and M. N. Sulaiman. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2):1–11, 2015. 2
- [13] Nathalie Japkowicz and Mohak Shah. *Evaluating Learning Algorithms: A Classification Perspective*. Cambridge University Press, 2011. 2
- [14] Pierre-Marc Jodoin, Sébastien Piérard, Yi Wang, and Marc Van Droogenbroeck. Overview and benchmarking of motion detection methods. In *Background Modeling and Foreground Detection for Video Surveillance*, chapter 24. Chapman and Hall/CRC, 2014. 7
- [15] Hermann Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30(5):509–541, 1977. 4, 13
- [16] Maurice George Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81, 1938. 3, 4
- [17] Andrey Nikolaevich Kolmogorov. Sur la notion de la moyenne. *Atti della Accademia Nazionale dei Lincei*, 12(9):388–391, 1930. 4
- [18] Yangguang Liu, Yangming Zhou, Shiting Wen, and Chao-gang Tang. A strategy on selecting performance metrics for classifier evaluation. *International Journal of Mobile Computing and Multimedia Communications*, 6(4):20–35, 2014. 2
- [19] Sébastien Piérard, Anaïs Halin, Anthony Cioppa, Adrien Deliège, and Marc Van Droogenbroeck. The Tile: A 2D map of ranking scores for two-class classification. *arXiv*, abs/2412.04309, 2024. 2
- [20] Sébastien Piérard, Anaïs Halin, Anthony Cioppa, Adrien Deliège, and Marc Van Droogenbroeck. Foundations of the theory of performance-based ranking. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14293–14302, Nashville, Tennessee, USA, 2025. IEEE. 2, 3, 5, 11, 12
- [21] David M. W. Powers. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1):37–63, 2011. 2
- [22] Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427–437, 2009. 2
- [23] Charles Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904. 3
- [24] Yi Wang, Pierre-Marc Jodoin, Fatih Porikli, Janusz Konrad, Yannick Benezeth, and Prakash Ishwar. CDnet 2014: An expanded change detection benchmark dataset. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 393–400, Columbus, Ohio, USA, 2014. Institute of Electrical and Electronics Engineers (IEEE). 7

A. Supplementary Material

Contents

A.1. Supplementary Material About Sec. 2.2.(Building Upon the Related Work)	11
A.1.1. All F_β Lead to Meaningful Performance Orderings	11
A.1.2. <i>SIVF</i> Leads to a Meaningful Performance Ordering When the Class Priors Are Fixed	11
A.1.3. <i>SIVF</i> Leads to the Same Ranking as F_β with $\beta^2 = \frac{\pi_-}{\pi_+}$	11
A.2 Supplementary Material About Sec. 3.(Theory)	13
A.2.1. On the Minimization of Fréchet Variance (Eq. (11))	13
A.2.2. On the Formulas for Degree of Optimality of Some F_β	13
A.3 Supplementary Material About Sec. 4.(Case Studies)	16
A.3.1. Principle of Analytical Computations for Kendall's rank correlations τ with Distributions of Performances	16
A.3.2. Detailed Results for the Uniform Distribution Over All Performances	17
A.3.3. Detailed Results for the Uniform Distributions With Fixed Probability of True Negatives	18
A.3.4. Detailed Results for the Uniform Distributions With Fixed Class Priors	20
A.3.5. Detailed Results for the Uniform Distributions With Fixed Class Priors, Above No-Skill	23
A.3.6. Detailed Results for the Uniform Distributions With Fixed Class Priors, Close to Oracle	26
A.3.7. All Results for Some Real Sets of Performances	29

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

A.1.1. All F_β Lead to Meaningful Performance Orderings

It has been shown in [20] that all performance orderings derived from some score of the form

$$R_I(P) = \frac{I(tn)P(\{tn\}) + I(tp)P(\{tp\})}{I(tn)P(\{tn\}) + I(fp)P(\{fp\}) + I(fn)P(\{fn\}) + I(tp)P(\{tp\})}, \quad (31)$$

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 [20], 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:

$$F_\beta : \{P \in \mathbb{P}_{(\Omega, \Sigma)} : P(\{tp\}) \neq 0\} \rightarrow [0, 1] : P \mapsto \left(\frac{1}{1 + \beta^2} Pr^{-1}(P) + \frac{\beta^2}{1 + \beta^2} Re^{-1}(P) \right)^{-1}. \quad (32)$$

The second one is

$$F_\beta : \{P \in \mathbb{P}_{(\Omega, \Sigma)} : P(\{tn\}) \neq 1\} \rightarrow [0, 1] : P \mapsto \frac{(1 + \beta^2) P(\{tp\})}{1 P(\{fp\}) + \beta^2 P(\{fn\}) + (1 + \beta^2) P(\{tp\})}. \quad (33)$$

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 $\{P \in \mathbb{P}_{(\Omega, \Sigma)} : P(\{tp\}) \neq 0\}$. It is now possible to compare Eq. (33) with Eq. (31) and to see that, for any given β , F_β is a particular case of ranking score with

$$(I(tn), I(fp), I(fn), I(tp)) \propto (0, 1, \beta^2, 1 + \beta^2), \quad (34)$$

or equivalently

$$(I(tn), I(fp), I(fn), I(tp)) \propto (0, 1 - b, b, 1), \quad (35)$$

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

A.1.2. SIVF Leads to a Meaningful Performance Ordering When the Class Priors Are Fixed

The skew-insensitive version of F_1 as been defined in [7] as

$$SIVF = \frac{2TPR}{TPR + FPR + 1}. \quad (36)$$

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

$$\begin{aligned} SIVF(P) &= \frac{2 \frac{P(\{tp\})}{\pi_+}}{\frac{P(\{tp\})}{\pi_+} + \frac{P(\{fp\})}{\pi_-} + 1} = \frac{2 \pi_- P(\{tp\})}{\pi_- P(\{tp\}) + \pi_+ P(\{fp\}) + \pi_- \pi_+} \\ &= \frac{2 \pi_- P(\{tp\})}{\pi_- P(\{tp\}) + \pi_+ P(\{fp\}) + \pi_- (P(\{fn\}) + P(\{tp\}))} = \frac{2 \pi_- P(\{tp\})}{\pi_+ P(\{fp\}) + \pi_- P(\{fn\}) + 2 \pi_- P(\{tp\})}. \end{aligned} \quad (37)$$

The comparison of Eq. (37) with Eq. (31) shows that $SIVF$ is a particular case of ranking score with

$$(I(tn), I(fp), I(fn), I(tp)) \propto (0, \pi_+, \pi_-, 2\pi_-). \quad (38)$$

A.1.3. 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 [7], 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

$$\beta^2 = \frac{\pi_-}{\pi_+} \Rightarrow F_\beta = \frac{SIVF}{(\pi_+ - \pi_-) SIVF + 2\pi_-}. \quad (39)$$

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

$$\beta^2 = \frac{\pi_-}{\pi_+} \Rightarrow \frac{\partial F_\beta}{\partial SIVF} > 0. \quad (40)$$

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 [20], 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. (34) and Eq. (38) shows that:

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

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. (11))

In Sec. 3.4.1, we say that the optimal tradeoffs are the F_β scores that minimize the Fréchet variance [8]:

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

We also say that the solutions, known as the Karcher means [15] are those that are equidistant of Pr and Re , i.e. such that

$$d_\tau(Pr; F_*) = d_\tau(F_*; Re) = \frac{d_\tau(Pr; Re)}{2}$$

$$\Leftrightarrow \tau(Pr; F_*) = \tau(F_*; Re) = \frac{1 + \tau(Pr; Re)}{2}.$$

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. (9) states that

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

we have

$$\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + (d_\tau(Pr; Re) - d_\tau(Pr; F_\beta))^2 \quad (41)$$

$$= 2d_\tau^2(Pr; F_\beta) - 2d_\tau(Pr; Re)d_\tau(Pr; F_\beta) + d_\tau^2(Pr; Re) \quad (42)$$

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

$$d_\tau(Pr; F_\beta) = \frac{d_\tau(Pr; Re)}{2}. \quad (43)$$

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 Formulas for Degree of Optimality of Some F_β

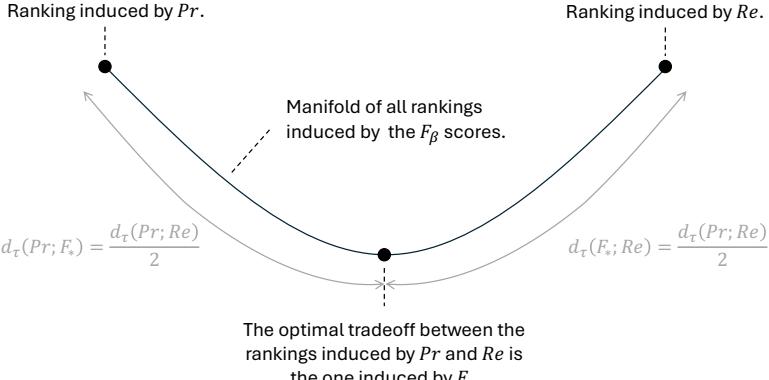
We provide here a few explanations about how Eqs. (16) to (18) can be derived. For that, let us remind three interpretations for Kendall's distance $d_\tau(X_1; X_2) \in [0, 1]$.

- First, Eq. (8) 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.
- 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 $d_\tau = \frac{1-\tau}{2}$.

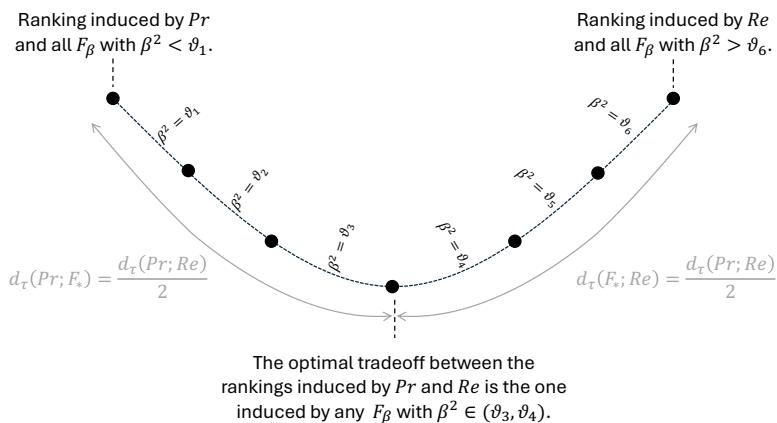
Keeping these three interpretations in mind is helpful to derive Eqs. (16) to (18).

Let us consider Eq. (17), for example. We are interested in the probability $P(\mathbf{x})$. 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.

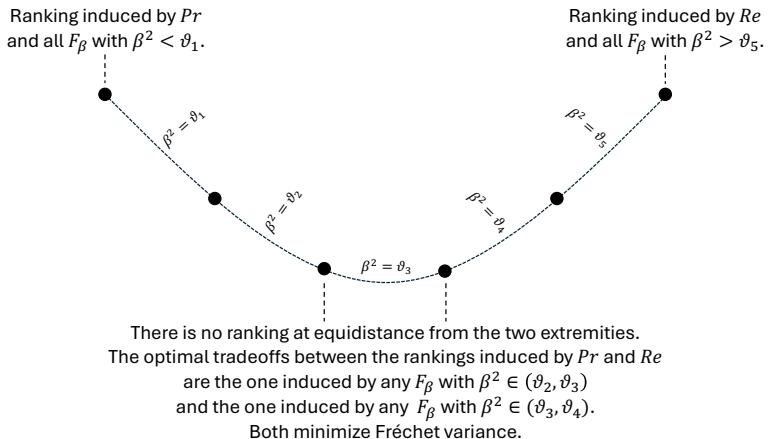
As shown in Fig. A.2.2, $d_\tau(F_\beta; F_*) = 1/2 |d_\tau(F_\beta; Re) - d_\tau(Pr; F_\beta)|$. Finally, using the third interpretation, we find that $1/2 |d_\tau(F_\beta; Re) - d_\tau(Pr; F_\beta)| = 1/4 |\tau(Pr; F_\beta) - \tau(F_\beta; Re)|$.



(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 $d_\tau(Pr; F_*) = d_\tau(F_*; Re)$.

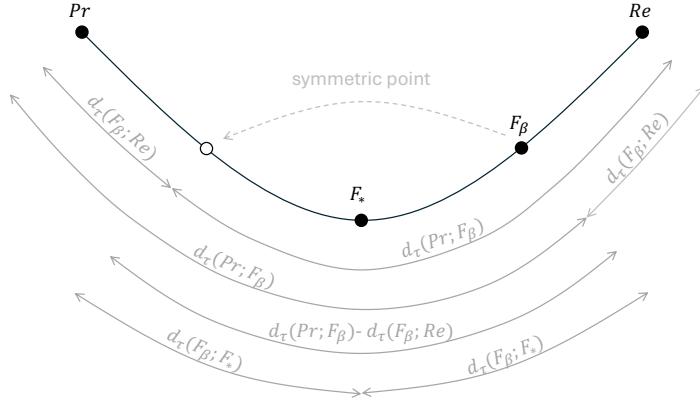


(b) Case in which the number of rankings induced by the F_β scores is odd. The rankings form a path graph. The nodes correspond to the rankings, each of them being induced by a range of values for β . The edges correspond to the swaps that occur between the consecutive rankings, at some given values 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 β s for which the ranking is at the midpoint where $d_\tau(Pr; F_*) = d_\tau(F_*; Re)$.

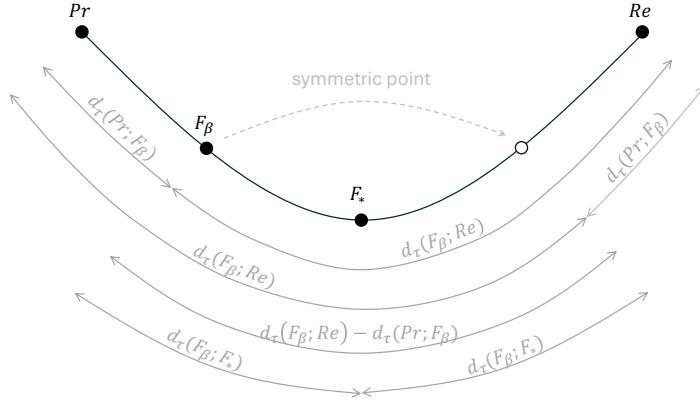


(c) Case in which the number of rankings induced by the F_β scores is even. Assuming there is no group of at least three co-aligned performances, the median of all the ϑ values corresponds to the swap between the two rankings minimizing Fréchet variance.

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



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

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. (8). Instead, we can compute analytically the following probabilities:

$$\text{Proba} [X_1(P_A) < X_1(P_B), X_2(P_A) < X_2(P_B)] = \int_{P_A \in \mathbb{P}_{(\Omega, \Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega, \Sigma)}} \mathbf{1}_{X_1(P_A) < X_1(P_B), X_2(P_A) < X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) dP_A dP_B, \quad (44)$$

$$\text{Proba} [X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)] = \int_{P_A \in \mathbb{P}_{(\Omega, \Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega, \Sigma)}} \mathbf{1}_{X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) dP_A dP_B, \quad (45)$$

$$\text{Proba} [X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)] = \int_{P_A \in \mathbb{P}_{(\Omega, \Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega, \Sigma)}} \mathbf{1}_{X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) dP_A dP_B, \quad (46)$$

$$\text{Proba} [X_1(P_A) > X_1(P_B), X_2(P_A) > X_2(P_B)] = \int_{P_A \in \mathbb{P}_{(\Omega, \Sigma)}} \int_{P_B \in \mathbb{P}_{(\Omega, \Sigma)}} \mathbf{1}_{X_1(P_A) > X_1(P_B), X_2(P_A) > X_2(P_B)} f_{\mathcal{P}}(P_A) f_{\mathcal{P}}(P_B) dP_A dP_B, \quad (47)$$

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

$$\tau(X_1; X_2) = 1 - 2 (\text{Proba} [X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)] + \text{Proba} [X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)]). \quad (48)$$

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

$$\text{Proba} [X_1(P_A) < X_1(P_B), X_2(P_A) < X_2(P_B)] = \text{Proba} [X_1(P_A) > X_1(P_B), X_2(P_A) > X_2(P_B)], \quad (49)$$

and

$$\text{Proba} [X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)] = \text{Proba} [X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)]. \quad (50)$$

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:

$$\tau(X_1; X_2) = 4 \text{Proba} [X_1(P_A) < X_1(P_B), X_2(P_A) < X_2(P_B)] - 1, \quad (51)$$

$$\tau(X_1; X_2) = 1 - 4 \text{Proba} [X_1(P_A) < X_1(P_B), X_2(P_A) > X_2(P_B)], \quad (52)$$

$$\tau(X_1; X_2) = 1 - 4 \text{Proba} [X_1(P_A) > X_1(P_B), X_2(P_A) < X_2(P_B)], \quad (53)$$

$$\tau(X_1; X_2) = 4 \text{Proba} [X_1(P_A) > X_1(P_B), X_2(P_A) > X_2(P_B)] - 1. \quad (54)$$

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.

```

1 Tetra = Simplex[{{1, 0, 0, 0}, {0, 1, 0, 0}, {0, 0, 1, 0}, {0, 0, 0, 1}}]
2 Ppv[ptn_, pfp_, pfn_, ptp_] := ptp/(pfp + ptp)
3 Tpr[ptn_, pfp_, pfn_, ptp_] := ptp/(pfn + ptp)
4 Fone[ptn_, pfp_, pfn_, ptp_] := 2*ptp/(pfp + pfn + 2*ptp)
5
6 tot = Integrate[
7   1,
8   {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
9   {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
10 ]

```

We have

$$\tau(Pr; Re) = \frac{1}{3}, \quad (55)$$

as shown with the code:

```

1 pLtLt = Integrate[
2   Boole[
3     Ppv[ptn1, pfp1, pfn1, ptp1] < Ppv[ptn2, pfp2, pfn2, ptp2] &&
4     Tpr[ptn1, pfp1, pfn1, ptp1] < Tpr[ptn2, pfp2, pfn2, ptp2]
5     ],
6     {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
7     {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
8   ] / tot
9 tau = 4 * pLtLt - 1

```

We have

$$\tau(Pr; F_1) = \frac{2}{3}, \quad (56)$$

as shown with the code:

```

1 pLtGt = Integrate[
2   Boole[
3     Ppv[ptn1, pfp1, pfn1, ptp1] < Ppv[ptn2, pfp2, pfn2, ptp2] &&
4     Fone[ptn1, pfp1, pfn1, ptp1] > Fone[ptn2, pfp2, pfn2, ptp2]
5     ],
6     {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
7     {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
8   ] / tot
9 tau = 1 - 4 * pLtGt

```

And we have

$$\tau(F_1; Re) = \frac{2}{3}, \quad (57)$$

as shown with the code:

```

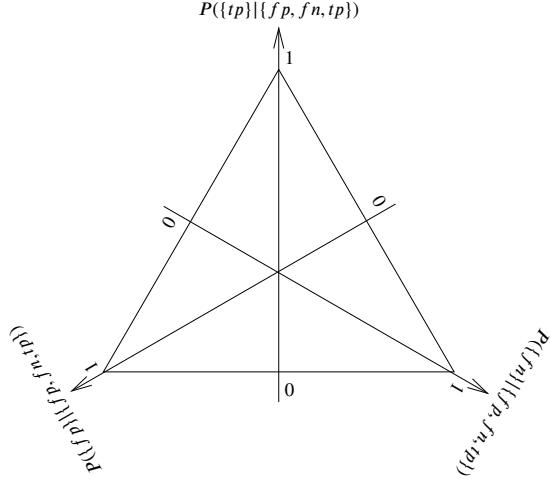
1 pLtGt = Integrate[
2   Boole[
3     Fone[ptn1, pfp1, pfn1, ptp1] < Fone[ptn2, pfp2, pfn2, ptp2] &&
4     Tpr[ptn1, pfp1, pfn1, ptp1] > Tpr[ptn2, pfp2, pfn2, ptp2]
5     ],
6     {ptn1, pfp1, pfn1, ptp1} \[Element] Tetra,
7     {ptn2, pfp2, pfn2, ptp2} \[Element] Tetra
8   ] / tot
9 tau = 1 - 4 * pLtGt

```

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.

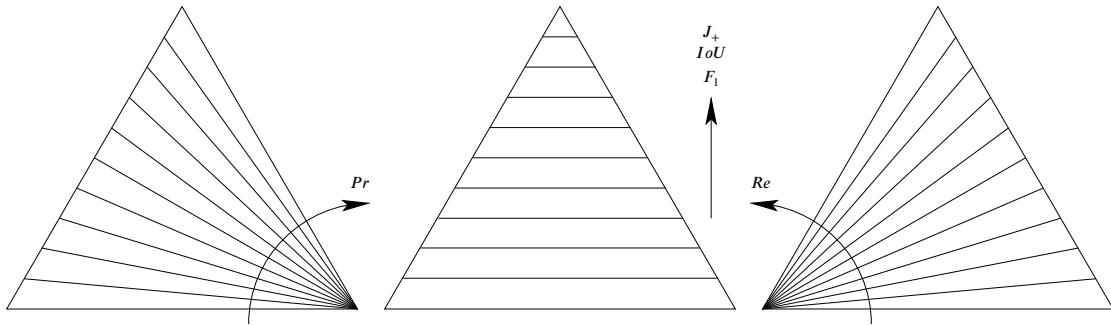
The equilateral triangles depicted in Fig. 3b form a "triangular space" in which the performances are assumed to be uniformly distributed. This "space" is parameterized as follows.



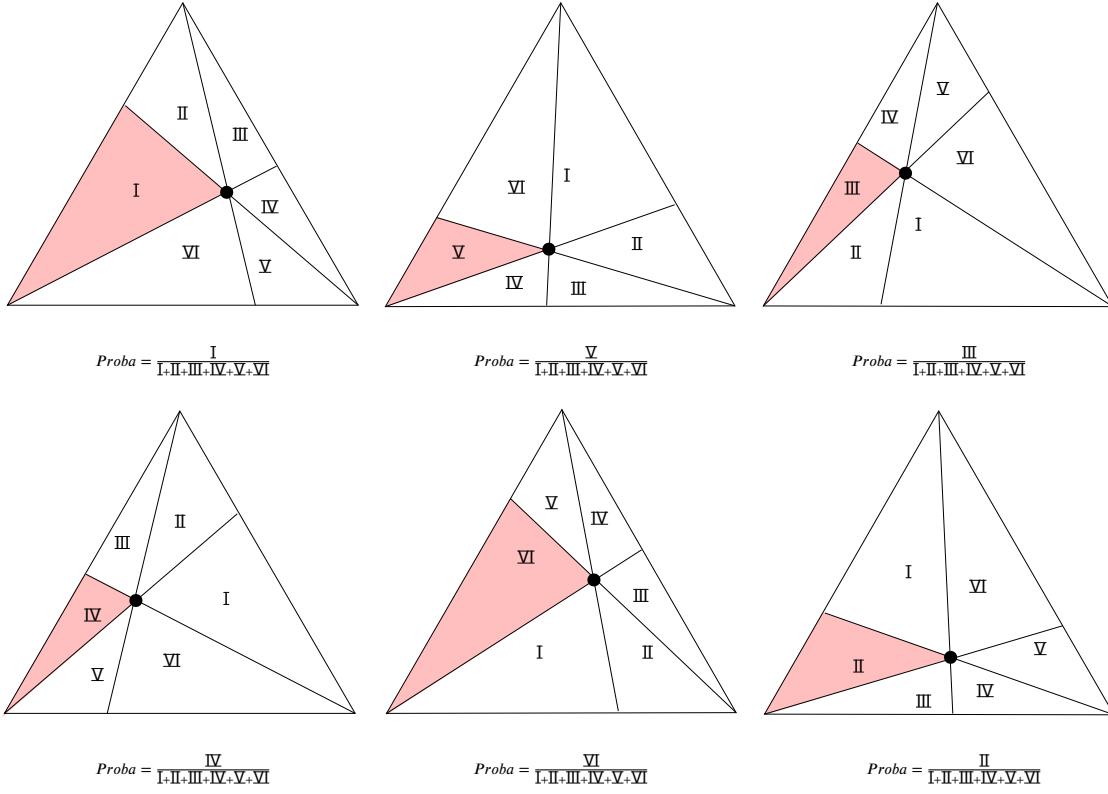
In fact, in the drawing shown here-above, the vertical axis corresponds to a score known as the *Intersection over Union IoU* or *Jaccard Coefficient* J_+ . It is monotonically increasing with F_1 :

$$F_1 = \frac{2IoU}{1 + IoU} \Rightarrow \frac{\partial F_1}{\partial IoU} > 0 \quad (58)$$

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.



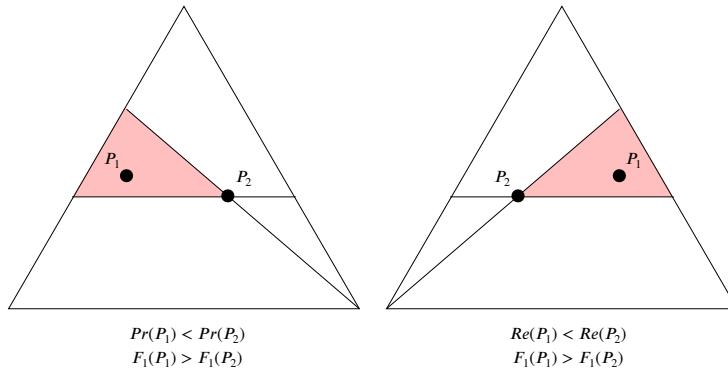
We now show that $\tau(Pr; Re) = 1/3$. In the following drawings, we consider a performance P_1 that can be anywhere in the "triangular space" and we choose one particular performance P_2 , shown as a black point. The pink area represents the probability that $Pr(P_1) < Pr(P_2)$ and $Re(P_1) > Re(P_2)$. 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 $\text{Pr}(P_1) < \text{Pr}(P_2)$ and $\text{Re}(P_1) > \text{Re}(P_2)$ is $1/6$. Using Eq. (52), we find the announced result:

$$\tau(\text{Pr}; \text{Re}) = 1/3 \quad (59)$$

We now show that $\tau(\text{Pr}; F_1) = \tau(\text{Re}; F_1)$. Using Eq. (52), we know that it is the case if and only if the probability that $\text{Pr}(P_1) < \text{Pr}(P_2)$ and $F_1(P_1) > F_1(P_2)$ is equal to the probability that $\text{Re}(P_1) < \text{Re}(P_2)$ 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. (52), we find the announced result:

$$\tau(\text{Pr}; F_1) = \tau(\text{Re}; F_1) \quad (60)$$

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

$$Pr(P_1) < Pr(P_2) \Leftrightarrow \frac{y_1}{x_1} < \frac{y_2}{x_2} \quad (61)$$

$$Re(P_1) < Re(P_2) \Leftrightarrow y_1 < y_2 \quad (62)$$

$$F_\beta(P_1) < F_\beta(P_2) \Leftrightarrow \frac{y_1}{x_1 + \ell} < \frac{y_2}{x_2 + \ell} \quad (63)$$

We start with a few initializations.

```

1 tot = Integrate[
2   1,
3   {x1, 0, 1}, {y1, 0, 1},
4   {x2, 0, 1}, {y2, 0, 1}
5 ]
```

We have

$$\tau(Pr; Re) = \frac{1}{2}, \quad (64)$$

as shown with the code (see Fig. A.3.3a):

```

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

We have

$$\tau(Pr; F_\beta) = 1 - \ell \left(\ell \log \left(\frac{\ell}{\ell+1} \right) + 1 \right), \quad (65)$$

as shown with the code (see Fig. A.3.3b):

```

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

And we have

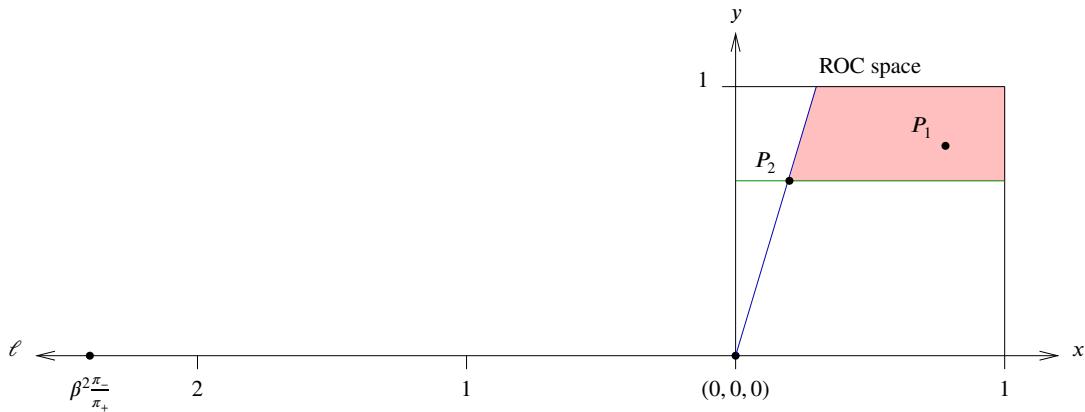
$$\tau(F_\beta; Re) = \ell^2 \left(-\log \left(\frac{1}{\ell} + 1 \right) \right) + \ell + \frac{1}{2}, \quad (66)$$

as shown with the code (see Fig. A.3.3c):

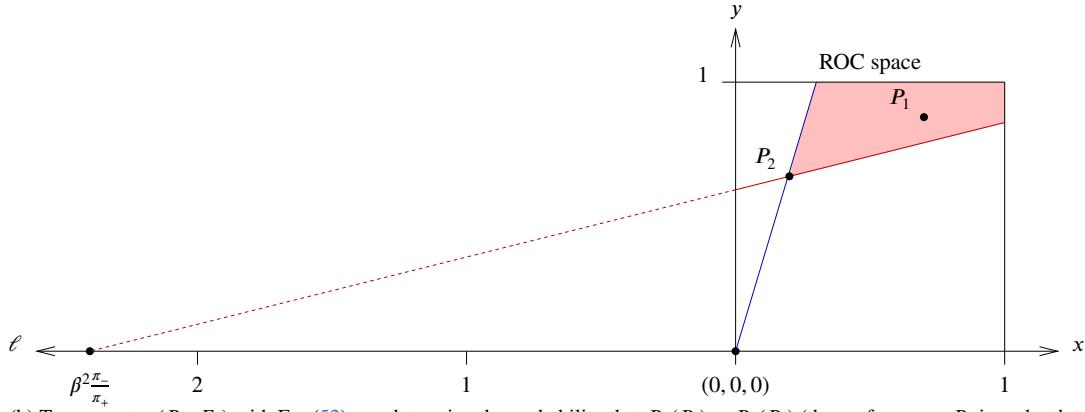
```

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

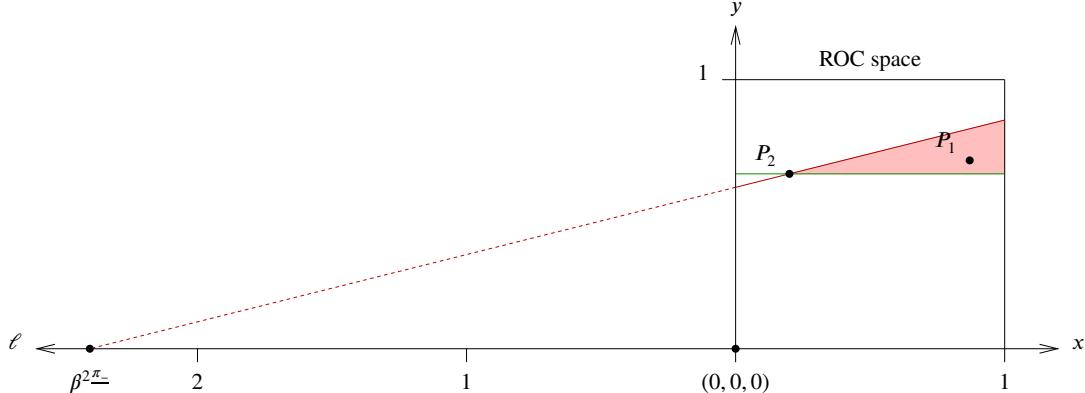
The plots are provided in Fig. A.3.4.



(a) To compute $\tau(Pr; Re)$ with Eq. (52), 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 .

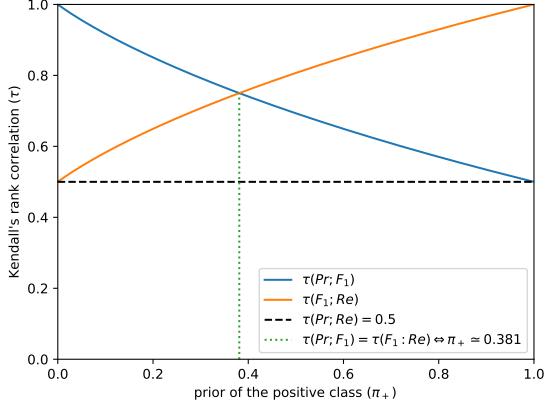


(b) To compute $\tau(Pr; F_\beta)$ with Eq. (52), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $F_\beta(P_1) > F_\beta(P_2)$ (the performance P_1 is above the red line), which is the pink area averaged for all positions of P_2 .

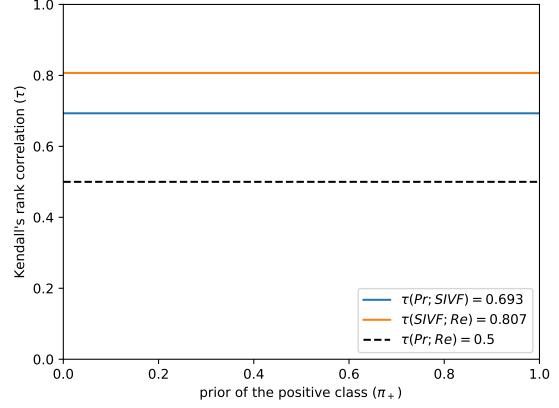


(c) To compute $\tau(F_\beta; Re)$ with Eq. (52), we determine the probability that $F_\beta(P_1) < F_\beta(P_2)$ (the performance P_1 is under the red line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2 .

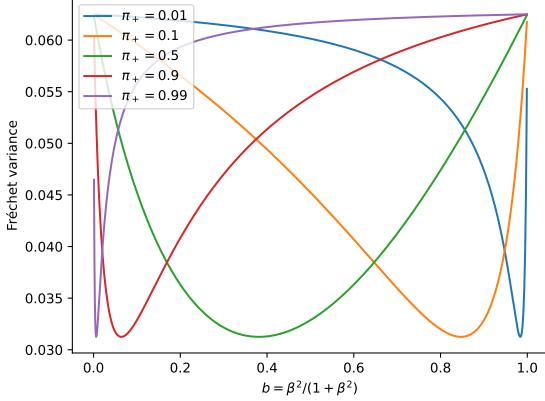
Figure A.3.3. Graphical representation, in and around the ROC space, of the principle that we use to derive the analytical expression for the optimal tradeoff between precision and recall for the purpose of ranking, in the case of uniform distributions with fixed class priors. Note that the ROC space is a linear mapping of the rectangles depicted in Fig. 3c.



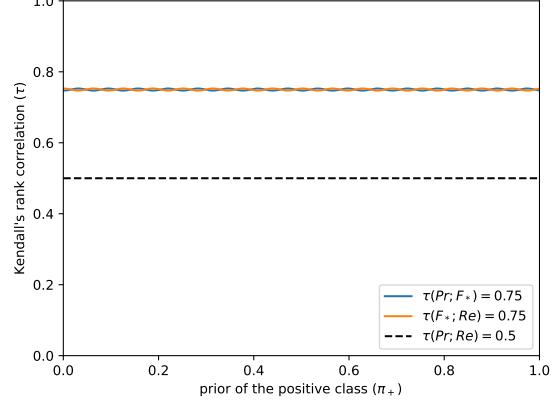
(a) F_1 , the traditional (balanced) F-score, is not the optimal tradeoff: $\tau(Pr; F_1) \neq \tau(F_1; Re)$.



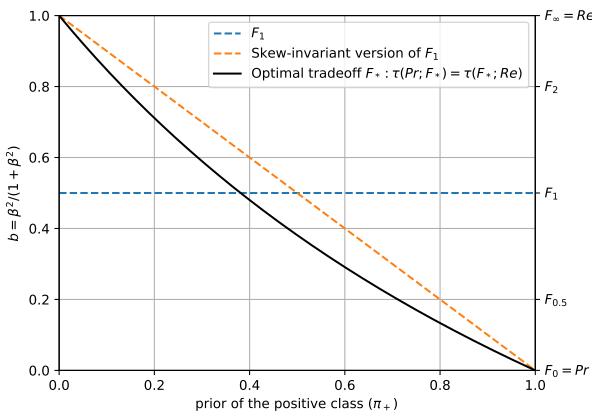
(b) $SIVF$, the skew-insensitive version of F_1 [7], is not the optimal tradeoff: $\tau(Pr; SIVF) \neq \tau(SIVF; Re)$.



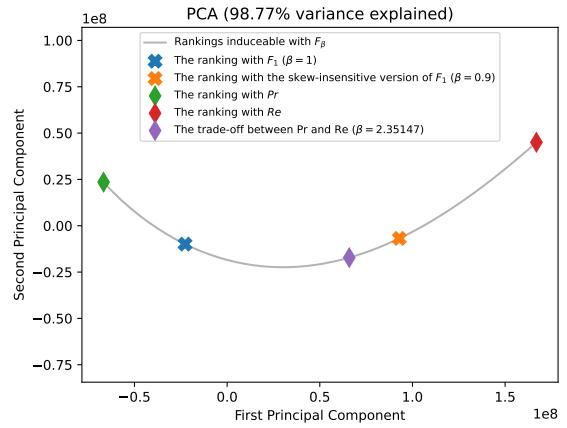
(c) Fréchet variance for various priors. It is defined at Eq. (11) and should be minimized to obtain the optimal tradeoff.



(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.3.4. 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

$$Pr(P_1) < Pr(P_2) \Leftrightarrow \frac{y_1}{x_1} < \frac{y_2}{x_2} \quad (67)$$

$$Re(P_1) < Re(P_2) \Leftrightarrow y_1 < y_2 \quad (68)$$

$$F_\beta(P_1) < F_\beta(P_2) \Leftrightarrow \frac{y_1}{x_1 + \ell} < \frac{y_2}{x_2 + \ell} \quad (69)$$

We start with a few initializations.

```

1 tot = Integrate[
2   1,
3   {x1, 0, 1}, {y1, x1, 1},
4   {x2, 0, 1}, {y2, x2, 1}
5 ]

```

We have

$$\tau(Pr; Re) = 0, \quad (70)$$

as shown with the code (see Fig. A.3.5a):

```

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

```

We have

$$\tau(Pr; F_\beta) = 1 - \frac{2}{3}\ell \left(-6\ell^2 - 6(\ell^2 - 1)\ell \log\left(\frac{\ell}{\ell + 1}\right) + 3\ell + 4 \right), \quad (71)$$

as shown with the code (see Fig. A.3.5b):

```

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

```

And we have

$$\tau(F_\beta; Re) = \frac{2}{3}\ell \left(-6\ell^2 + 6(\ell^2 - 1)\ell \log\left(\frac{1}{\ell} + 1\right) + 3\ell + 4 \right), \quad (72)$$

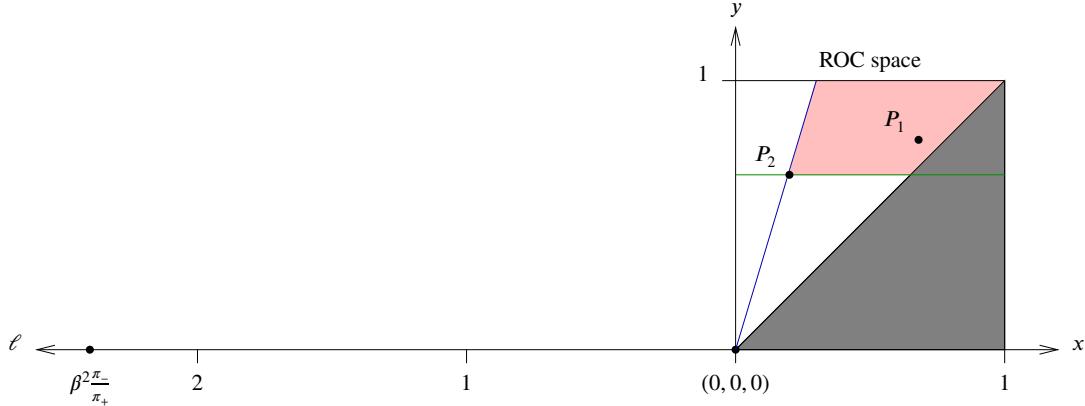
as shown with the code (see Fig. A.3.5c):

```

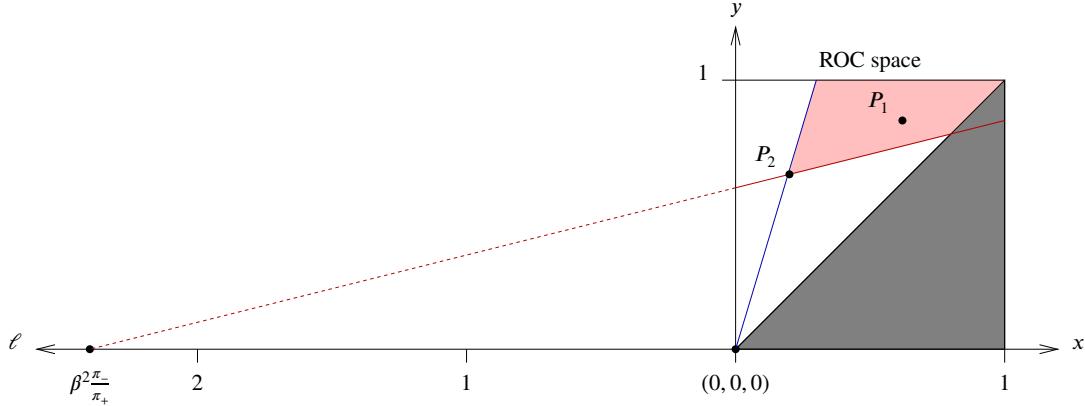
1 pLtGt = Integrate[
2   Boole[(y1 (x2 + 1) < y2 (x1 + 1)) && (y1 > y2)],
3   {x1, 0, 1}, {y1, x1, 1},
4   {x2, 0, 1}, {y2, x2, 1},
5   Assumptions -> 1 > 0
6 ] / tot
7 tau = 1 - 4 * pLtGt
8 tau = FullSimplify[tau, Assumptions -> 1 > 0]

```

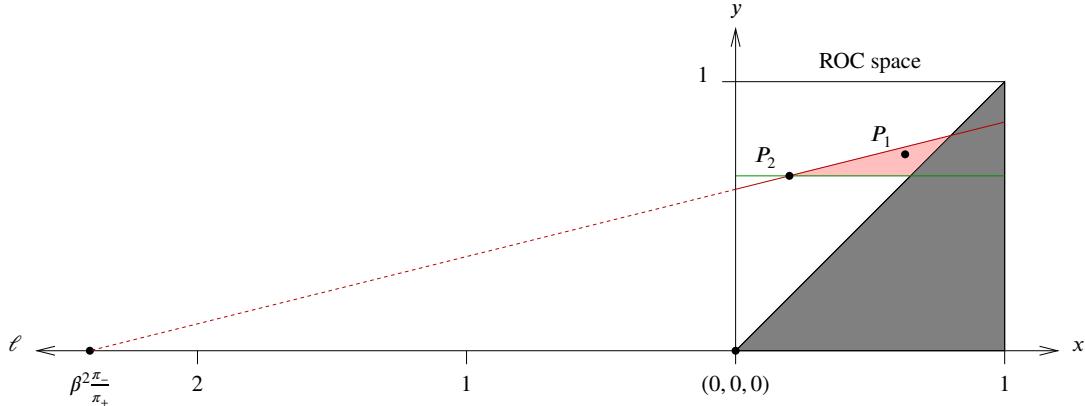
The plots are provided in Fig. A.3.6.



(a) To compute $\tau(Pr; Re)$ with Eq. (52), 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 .

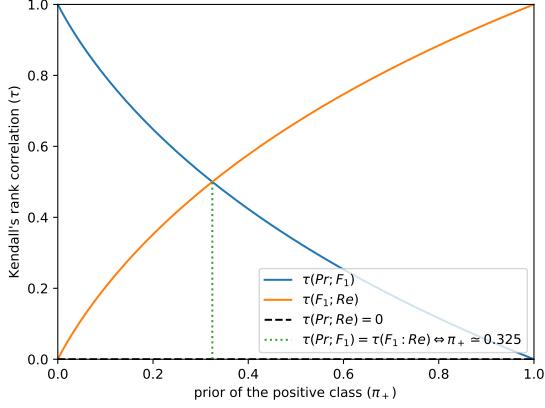


(b) To compute $\tau(Pr; F_\beta)$ with Eq. (52), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $F_\beta(P_1) > F_\beta(P_2)$ (the performance P_1 is above the red line), which is the pink area averaged for all positions of P_2 .

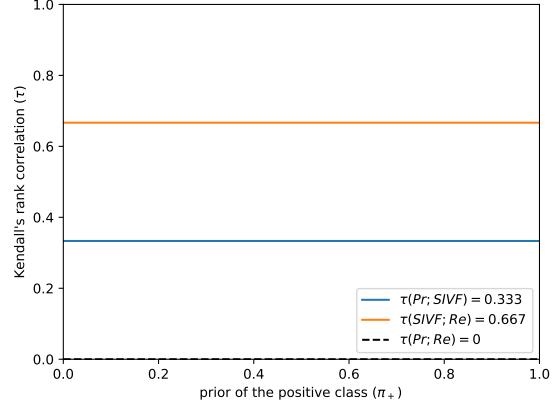


(c) To compute $\tau(F_\beta; Re)$ with Eq. (52), we determine the probability that $F_\beta(P_1) < F_\beta(P_2)$ (the performance P_1 is under the red line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2 .

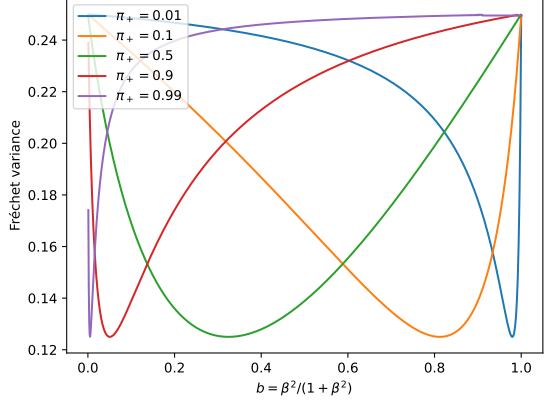
Figure A.3.5. Graphical representation, in and around the ROC space, of the principle that we use to derive the analytical expression for the optimal tradeoff between precision and recall for the purpose of ranking, in the case of uniform distributions with fixed class priors above no-skill. Note that the unshaded area of ROC space is a linear mapping of the triangles depicted in Fig. 3d.



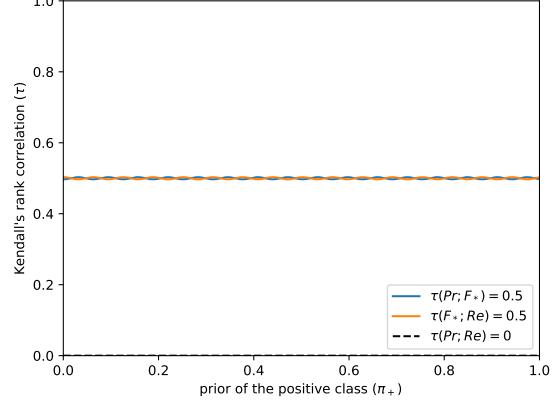
(a) F_1 , the traditional (balanced) F-score, is not the optimal tradeoff: $\tau(Pr; F_1) \neq \tau(F_1; Re)$.



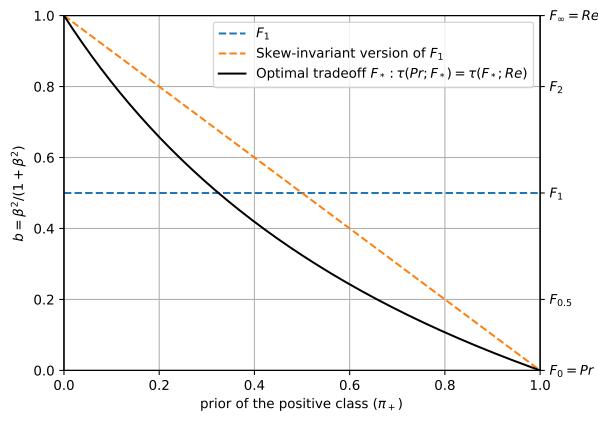
(b) $SIVF$, the skew-insensitive version of F_1 [7], is not the optimal tradeoff: $\tau(Pr; SIVF) \neq \tau(SIVF; Re)$.



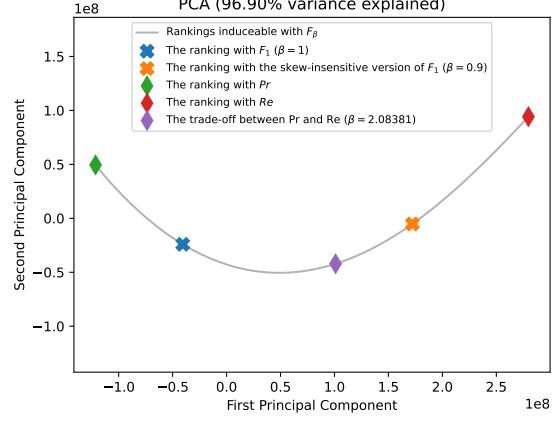
(c) Fréchet variance for various priors. It is defined at Eq. (11) and should be minimized to obtain the optimal tradeoff.



(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.3.6. 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

$$Pr(P_1) < Pr(P_2) \Leftrightarrow \frac{y_1}{x_1} < \frac{y_2}{x_2} \quad (73)$$

$$Re(P_1) < Re(P_2) \Leftrightarrow y_1 < y_2 \quad (74)$$

$$F_\beta(P_1) < F_\beta(P_2) \Leftrightarrow \frac{y_1}{x_1 + \ell} < \frac{y_2}{x_2 + \ell} \quad \text{with} \quad \ell = \beta^2 \frac{\pi_+}{\pi_-} = \frac{b}{1-b} \frac{\pi_+}{1-\pi_+} \quad (75)$$

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.

```

1 tot = Integrate[
2   1,
3   {x1, 0, p}, {y1, p, 1},
4   {x2, 0, p}, {y2, p, 1},
5   Assumptions -> 0 < p < 1
6 ]

```

We have

$$\tau(Pr; Re) = 1 - \frac{-\pi_+^4 + 2\pi_+^4 \log(\pi_+) + \pi_+^2}{2(1 - \pi_+)^2 \pi_+^2}, \quad (76)$$

as shown with the code (see Fig. A.3.7a):

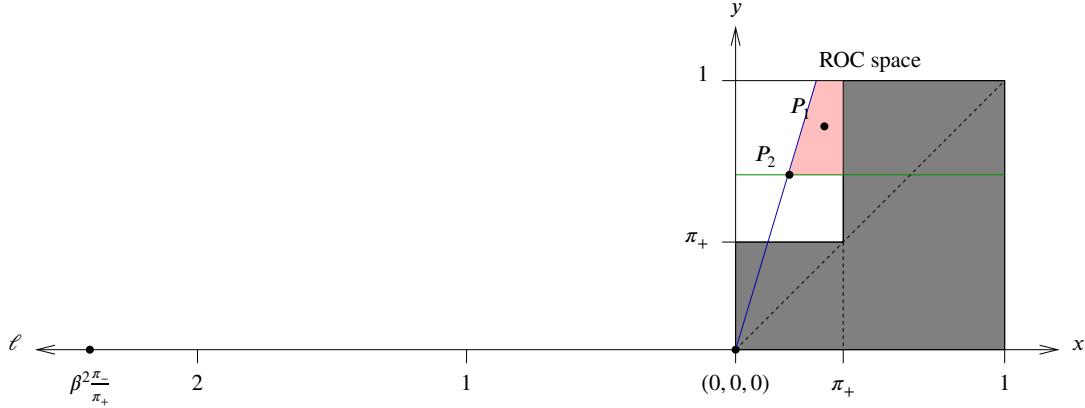
```

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

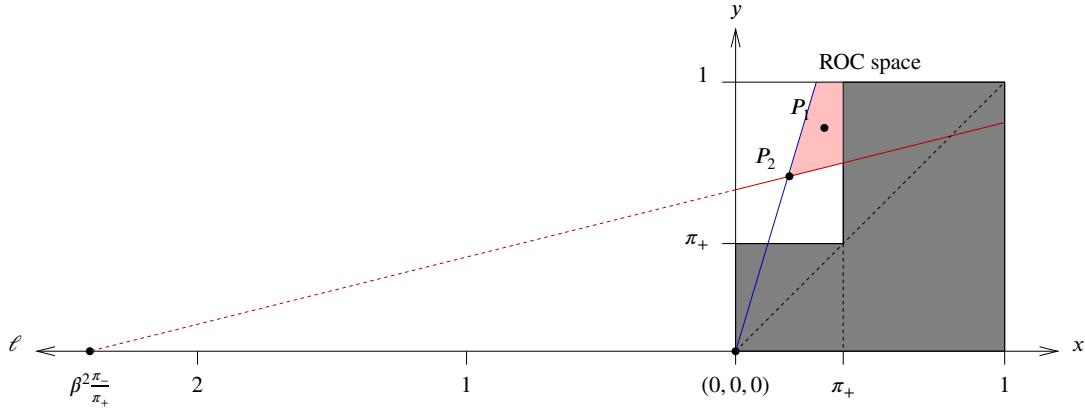
```

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.

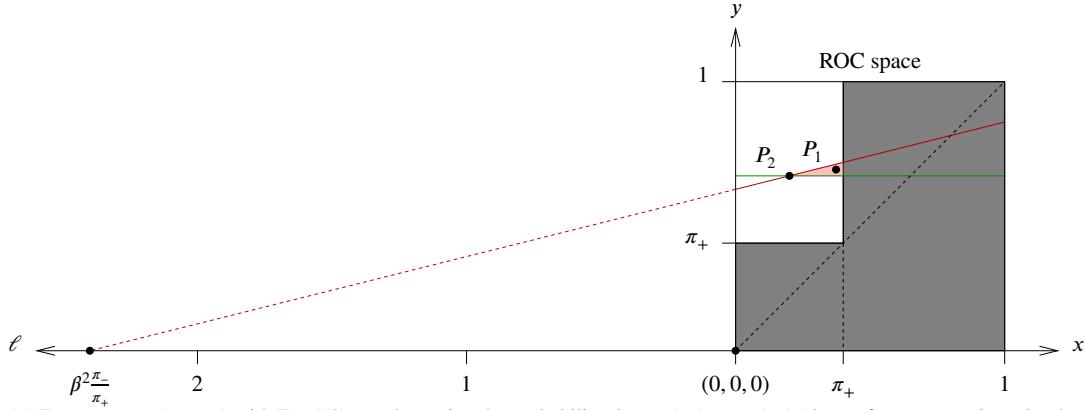
The plots are provided in Fig. A.3.8.



(a) To compute $\tau(Pr; Re)$ with Eq. (52), 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 .

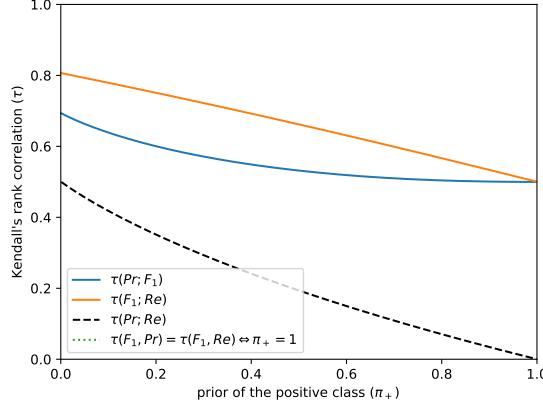


(b) To compute $\tau(Pr; F_\beta)$ with Eq. (52), we determine the probability that $Pr(P_1) < Pr(P_2)$ (the performance P_1 is under the blue line) and $F_\beta(P_1) > F_\beta(P_2)$ (the performance P_1 is above the red line), which is the pink area averaged for all positions of P_2 .

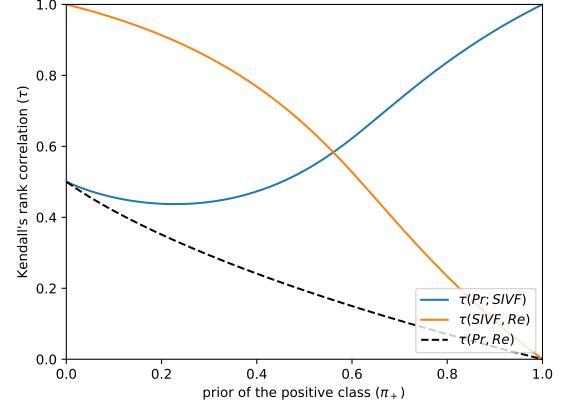


(c) To compute $\tau(F_\beta; Re)$ with Eq. (52), we determine the probability that $F_\beta(P_1) < F_\beta(P_2)$ (the performance P_1 is under the red line) and $Re(P_1) > Re(P_2)$ (the performance P_1 is above the green line), which is the pink area averaged for all positions of P_2 .

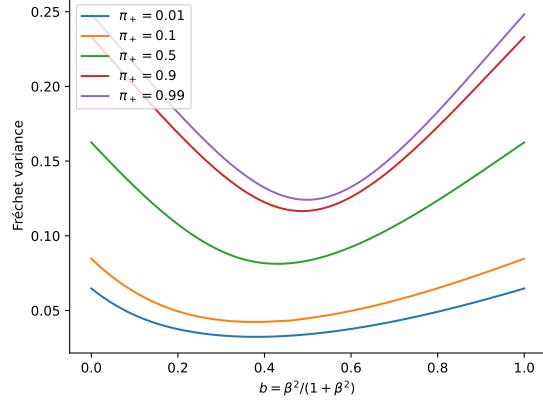
Figure A.3.7. Graphical representation, in and around the ROC space, of the principle that we use to derive the analytical expression for the optimal tradeoff between precision and recall for the purpose of ranking, in the case of uniform distributions with fixed class priors close to oracle. Note that the unshaded area of ROC space is a linear mapping of the squares depicted in Fig. 3e.



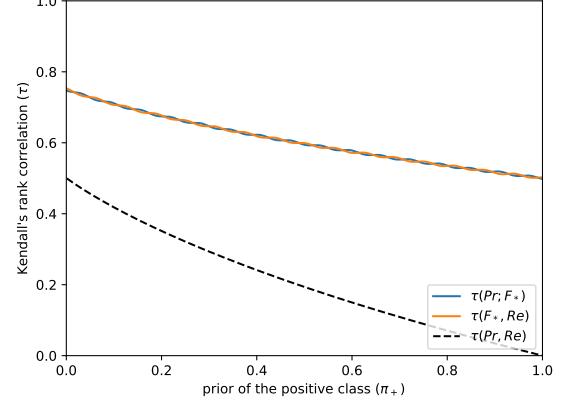
(a) F_1 , the traditional (balanced) F-score, is not the optimal tradeoff: $\tau(Pr; F_1) \neq \tau(F_1; Re)$.



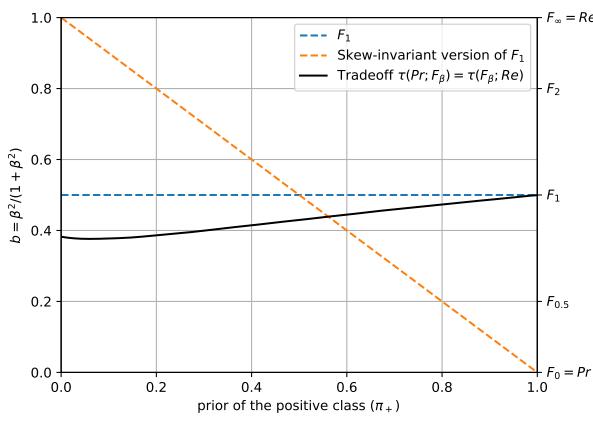
(b) $SIVF$, the skew-insensitive version of F_1 [7], is not the optimal tradeoff: $\tau(Pr; SIVF) \neq \tau(SIVF; Re)$.



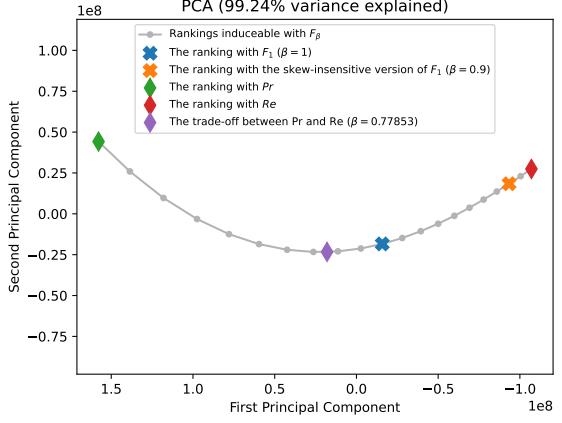
(c) Fréchet variance for various priors. It is defined at Eq. (11) and should be minimized to obtain the optimal tradeoff.



(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.3.8. 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.3.9.
- The results for video "PETS2006" (ranking of 57 methods; $\pi_+ = 0.0130$) can be found at Fig. A.3.10.
- The results for video "office" (ranking of 57 methods; $\pi_+ = 0.0690$) can be found at Fig. A.3.11.
- The results for video "highway" (ranking of 57 methods; $\pi_+ = 0.0593$) can be found at Fig. A.3.12.

Category "Dynamic Background"

- The results for video "overpass" (ranking of 57 methods; $\pi_+ = 0.0134$) can be found at Fig. A.3.13.
- The results for video "canoe" (ranking of 57 methods; $\pi_+ = 0.0354$) can be found at Fig. A.3.14.
- The results for video "fountain01" (ranking of 57 methods; $\pi_+ = 0.0008$) can be found at Fig. A.3.15.
- The results for video "fountain02" (ranking of 57 methods; $\pi_+ = 0.0022$) can be found at Fig. A.3.16.
- The results for video "fall" (ranking of 57 methods; $\pi_+ = 0.0177$) can be found at Fig. A.3.17.
- The results for video "boats" (ranking of 57 methods; $\pi_+ = 0.0063$) can be found at Fig. A.3.18.

Category "Camera Jitter"

- The results for video "boulevard" (ranking of 57 methods; $\pi_+ = 0.0469$) can be found at Fig. A.3.19.
- The results for video "sidewalk" (ranking of 57 methods; $\pi_+ = 0.0261$) can be found at Fig. A.3.20.
- The results for video "badminton" (ranking of 57 methods; $\pi_+ = 0.0343$) can be found at Fig. A.3.21.
- The results for video "traffic" (ranking of 57 methods; $\pi_+ = 0.0623$) can be found at Fig. A.3.22.

Category "Intermittent Object Motion"

- The results for video "abandonedBox" (ranking of 57 methods; $\pi_+ = 0.0481$) can be found at Fig. A.3.23.
- The results for video "winterDriveway" (ranking of 57 methods; $\pi_+ = 0.0075$) can be found at Fig. A.3.24.
- The results for video "sofa" (ranking of 57 methods; $\pi_+ = 0.0437$) can be found at Fig. A.3.25.
- The results for video "tramstop" (ranking of 57 methods; $\pi_+ = 0.1795$) can be found at Fig. A.3.26.
- The results for video "parking" (ranking of 57 methods; $\pi_+ = 0.0773$) can be found at Fig. A.3.27.
- The results for video "streetLight" (ranking of 57 methods; $\pi_+ = 0.0485$) can be found at Fig. A.3.28.

Category "Shadow"

- The results for video "copyMachine" (ranking of 57 methods; $\pi_+ = 0.0693$) can be found at Fig. A.3.29.
- The results for video "bungalows" (ranking of 57 methods; $\pi_+ = 0.0600$) can be found at Fig. A.3.30.
- The results for video "busStation" (ranking of 57 methods; $\pi_+ = 0.0369$) can be found at Fig. A.3.31.
- The results for video "peopleInShade" (ranking of 57 methods; $\pi_+ = 0.0564$) can be found at Fig. A.3.32.
- The results for video "backdoor" (ranking of 57 methods; $\pi_+ = 0.0199$) can be found at Fig. A.3.33.
- The results for video "cubicle" (ranking of 57 methods; $\pi_+ = 0.0196$) can be found at Fig. A.3.34.

Category "Thermal"

- The results for video "lakeSide" (ranking of 57 methods; $\pi_+ = 0.0192$) can be found at Fig. A.3.35.
- The results for video "diningRoom" (ranking of 57 methods; $\pi_+ = 0.0859$) can be found at Fig. A.3.36.
- The results for video "park" (ranking of 57 methods; $\pi_+ = 0.0203$) can be found at Fig. A.3.37.
- The results for video "corridor" (ranking of 57 methods; $\pi_+ = 0.0331$) can be found at Fig. A.3.38.
- The results for video "library" (ranking of 57 methods; $\pi_+ = 0.1928$) can be found at Fig. A.3.39.

Category "Bad Weather"

- The results for video "skating" (ranking of 57 methods; $\pi_+ = 0.0397$) can be found at Fig. A.3.40.
- The results for video "wetSnow" (ranking of 57 methods; $\pi_+ = 0.0150$) can be found at Fig. A.3.41.
- The results for video "snowFall" (ranking of 57 methods; $\pi_+ = 0.0105$) can be found at Fig. A.3.42.
- The results for video "blizzard" (ranking of 57 methods; $\pi_+ = 0.0115$) can be found at Fig. A.3.43.

Category "Low Framerate"

- The results for video "tunnelExit_0_35fps" (ranking of 57 methods; $\pi_+ = 0.0195$) can be found at Fig. A.3.44.
- The results for video "port_0_17fps" (ranking of 57 methods; $\pi_+ = 0.0002$) can be found at Fig. A.3.45.
- The results for video "tramCrossroad_1fps" (ranking of 57 methods; $\pi_+ = 0.0288$) can be found at Fig. A.3.46.
- The results for video "turnpike_0_5fps" (ranking of 57 methods; $\pi_+ = 0.0581$) can be found at Fig. A.3.47.

Category "Night Videos"

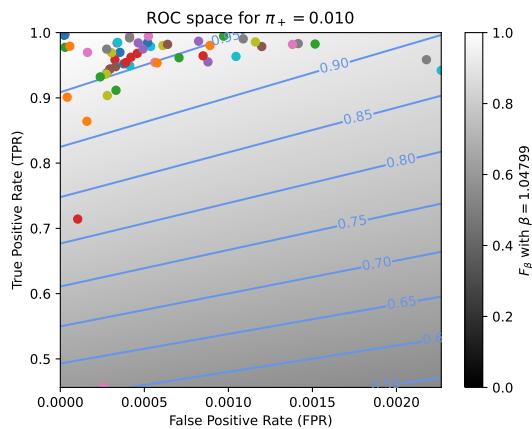
- The results for video "tramStation" (ranking of 58 methods; $\pi_+ = 0.0339$) can be found at Fig. A.3.48.
- The results for video "busyBoulvard" (ranking of 58 methods; $\pi_+ = 0.0847$) can be found at Fig. A.3.49.
- The results for video "streetCornerAtNight" (ranking of 58 methods; $\pi_+ = 0.0062$) can be found at Fig. A.3.50.
- The results for video "fluidHighway" (ranking of 58 methods; $\pi_+ = 0.0175$) can be found at Fig. A.3.51.
- The results for video "bridgeEntry" (ranking of 58 methods; $\pi_+ = 0.0200$) can be found at Fig. A.3.52.
- The results for video "winterStreet" (ranking of 58 methods; $\pi_+ = 0.0689$) can be found at Fig. A.3.53.

Category "PTZ"

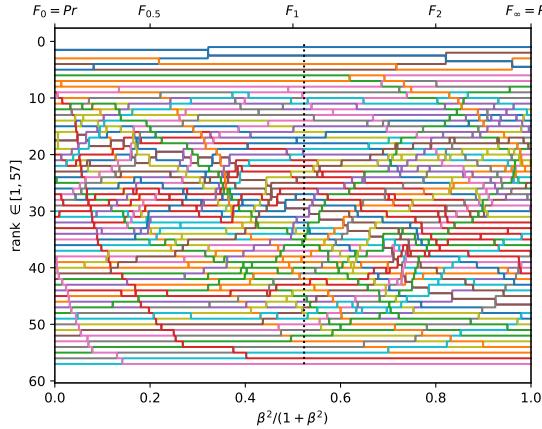
- The results for video "twoPositionPTZCam" (ranking of 56 methods; $\pi_+ = 0.0117$) can be found at Fig. A.3.54.
- The results for video "zoomInZoomOut" (ranking of 56 methods; $\pi_+ = 0.0019$) can be found at Fig. A.3.55.
- The results for video "continuousPan" (ranking of 56 methods; $\pi_+ = 0.0056$) can be found at Fig. A.3.56.
- The results for video "intermittentPan" (ranking of 56 methods; $\pi_+ = 0.0094$) can be found at Fig. A.3.57.

Category "Turbulence"

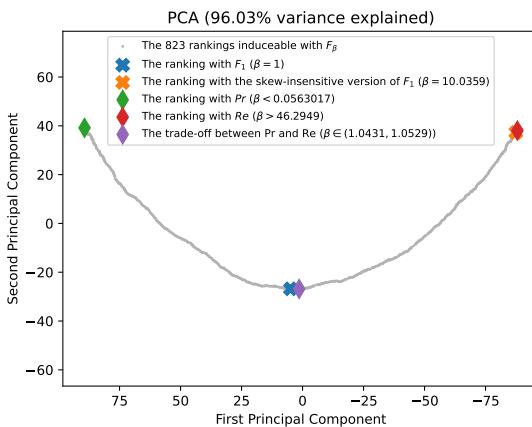
- The results for video "turbulence2" (ranking of 57 methods; $\pi_+ = 0.0008$) can be found at Fig. A.3.58.
- The results for video "turbulence3" (ranking of 57 methods; $\pi_+ = 0.0121$) can be found at Fig. A.3.59.
- The results for video "turbulence0" (ranking of 57 methods; $\pi_+ = 0.0015$) can be found at Fig. A.3.60.
- The results for video "turbulence1" (ranking of 57 methods; $\pi_+ = 0.0026$) can be found at Fig. A.3.61.



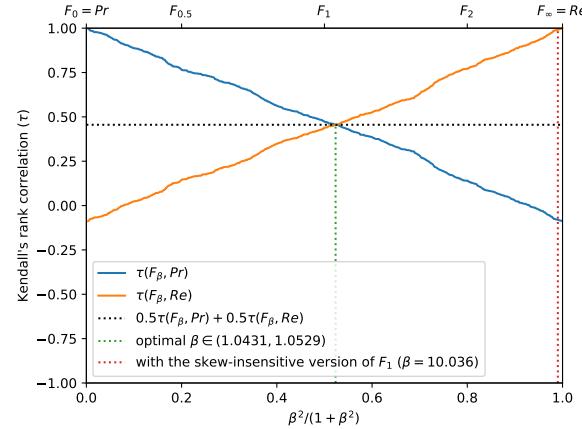
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



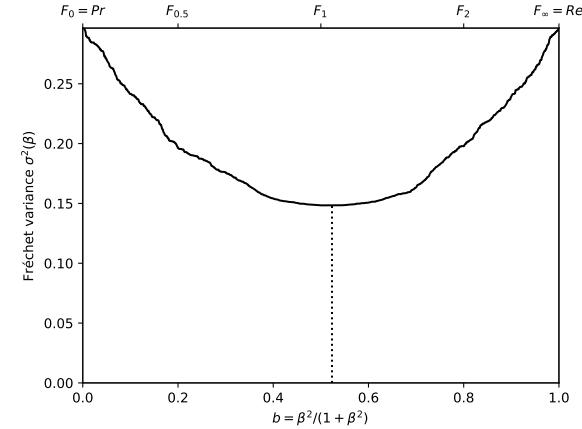
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



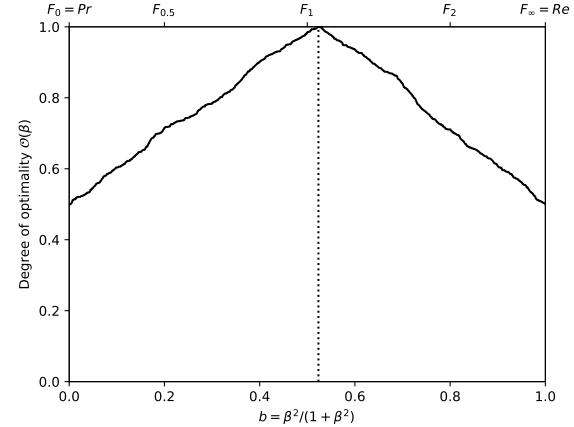
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

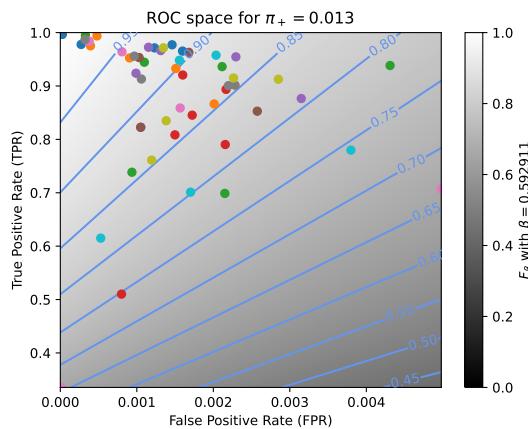


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

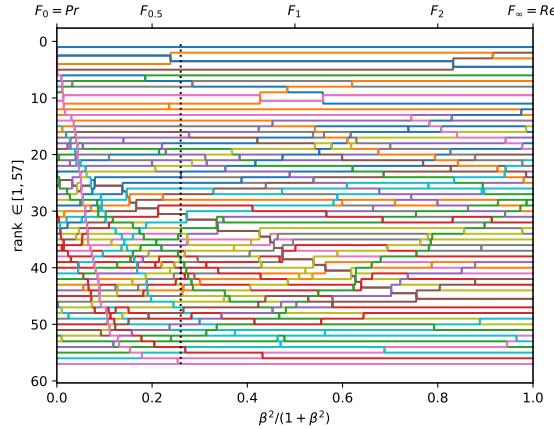


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

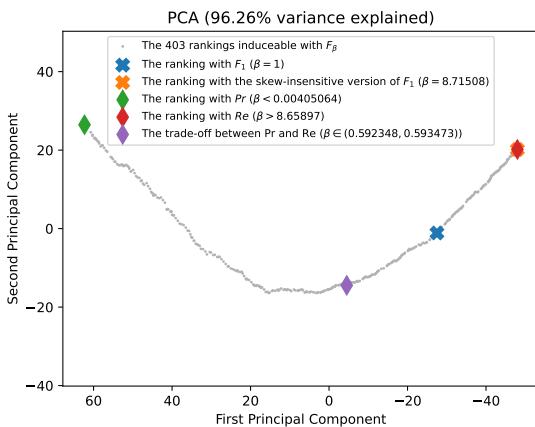
Figure A.3.9. Ranking of 57 BGS methods evaluated on the video "pedestrians" ($\pi_+ = 0.0098$).



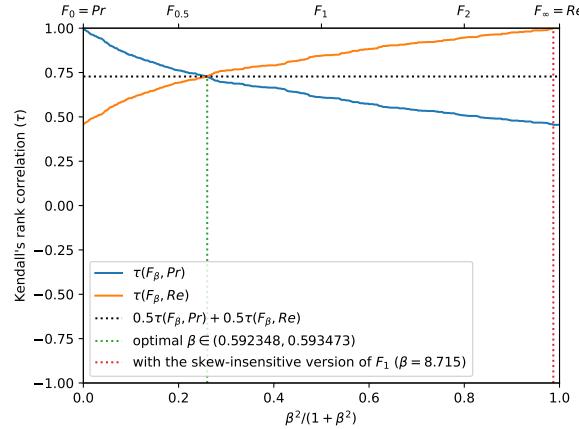
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



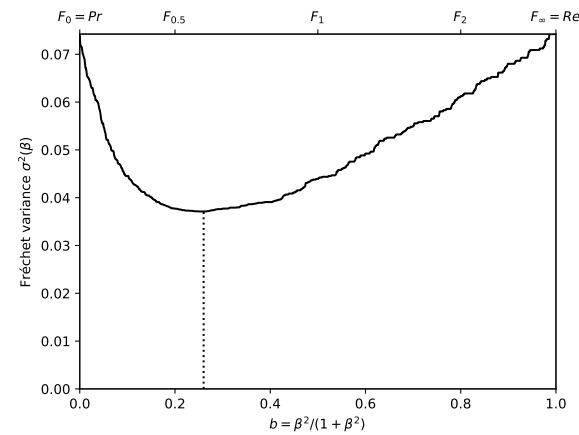
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



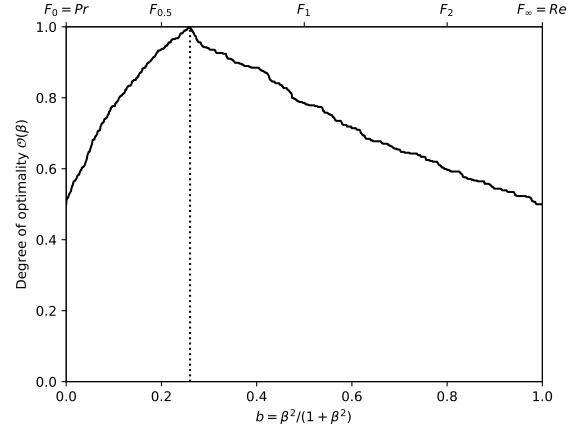
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

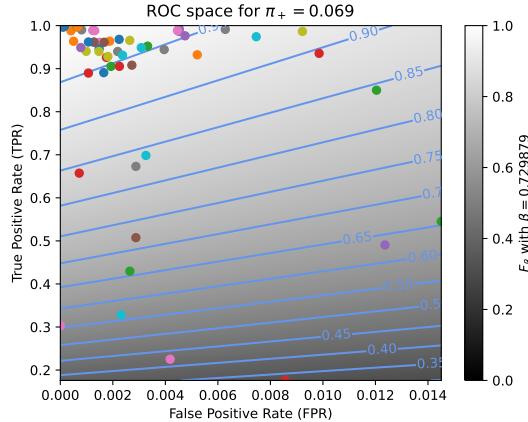


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

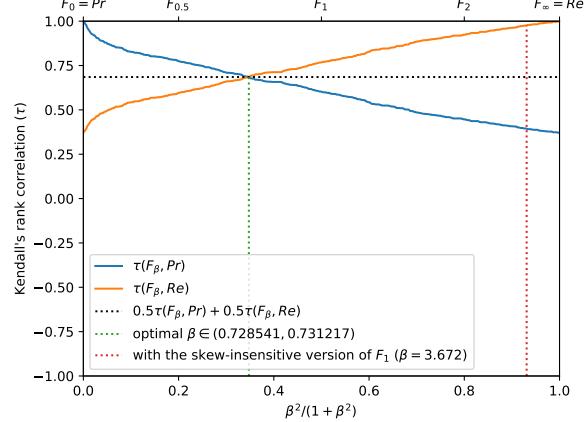


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

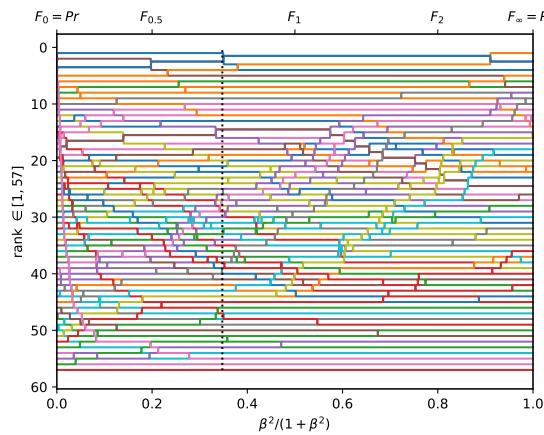
Figure A.3.10. Ranking of 57 BGS methods evaluated on the video "PETS2006" ($\pi_+ = 0.0130$).



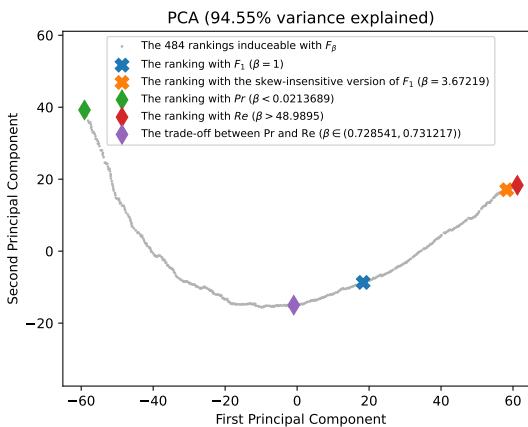
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



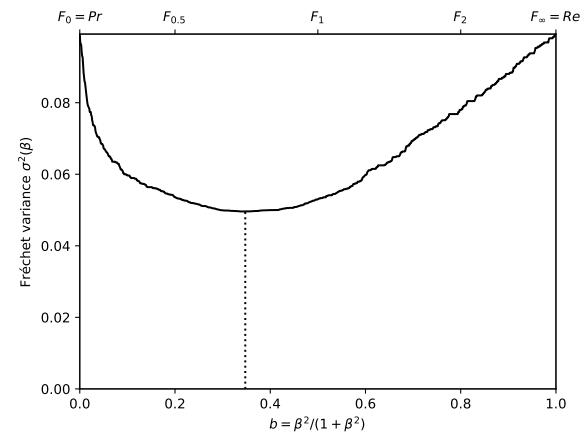
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



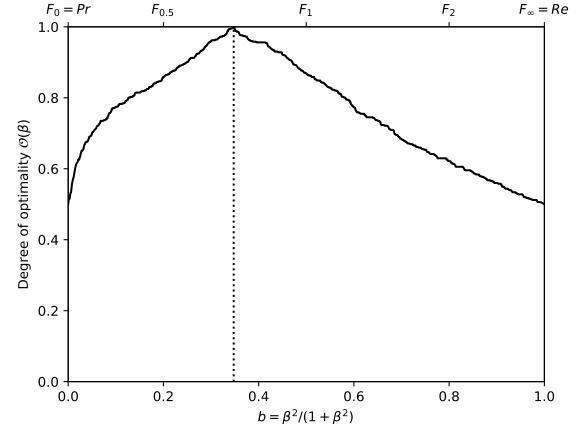
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

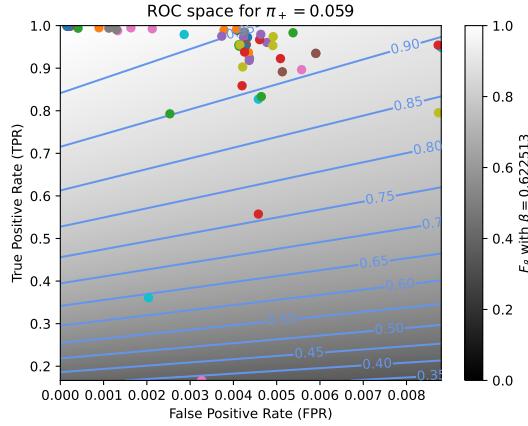


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

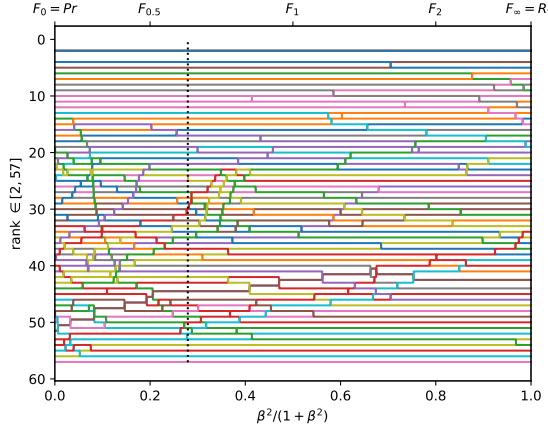


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

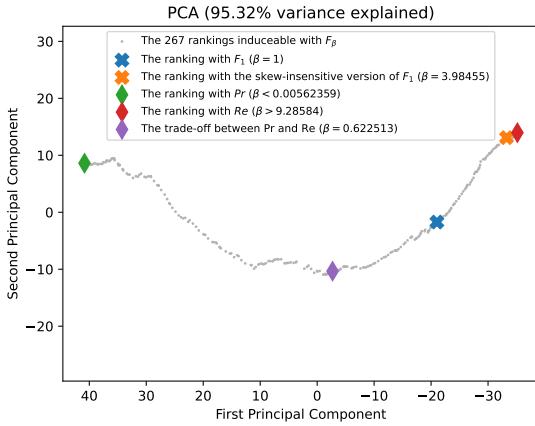
Figure A.3.11. Ranking of 57 BGS methods evaluated on the video "office" ($\pi_+ = 0.0690$).



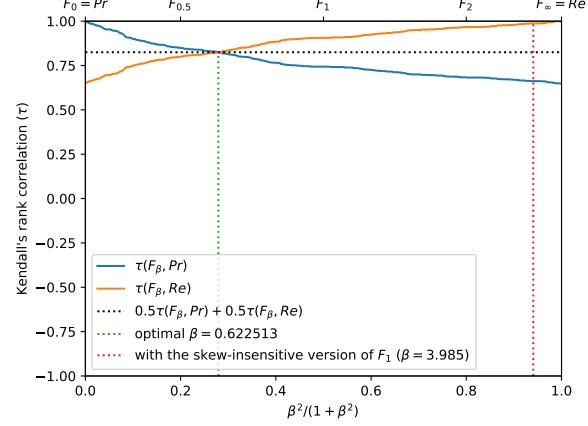
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



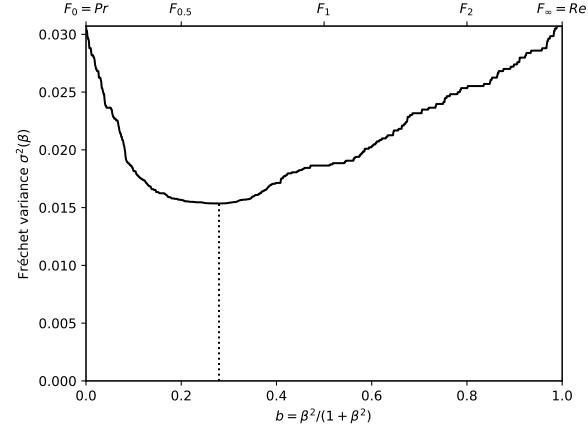
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



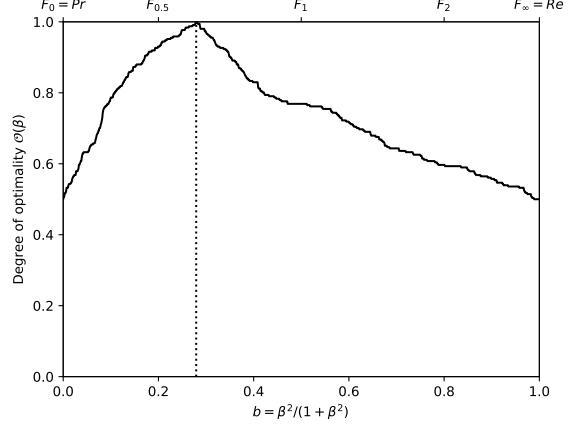
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

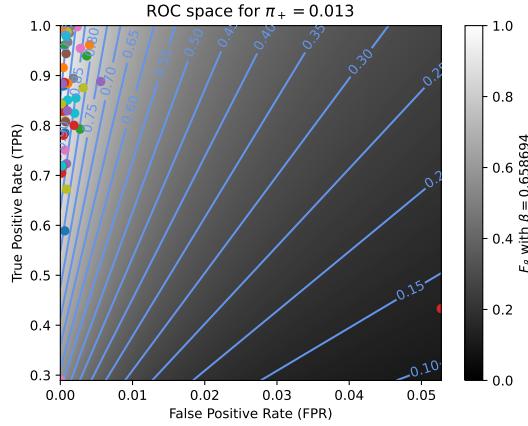


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

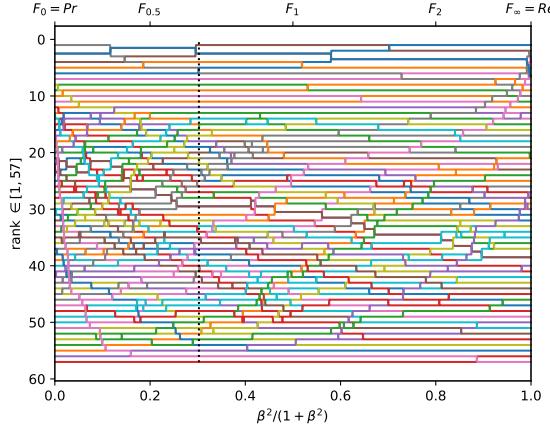


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

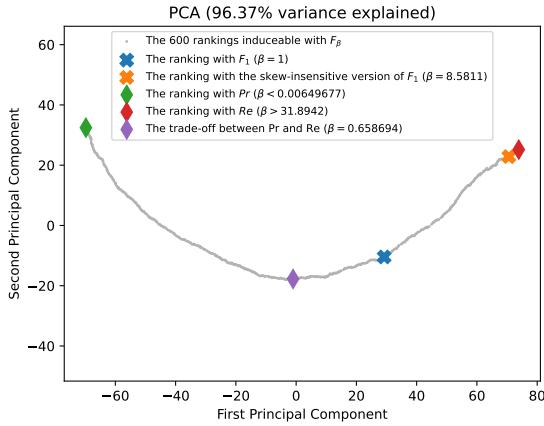
Figure A.3.12. Ranking of 57 BGS methods evaluated on the video "highway" ($\pi_+ = 0.0593$).



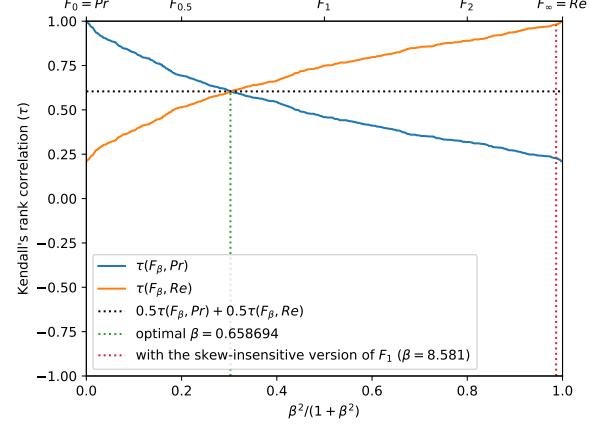
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



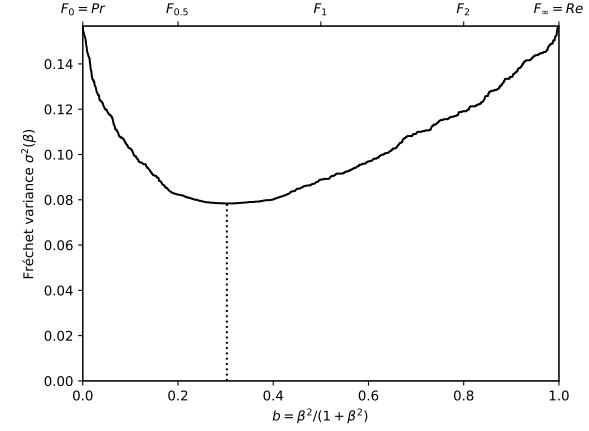
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



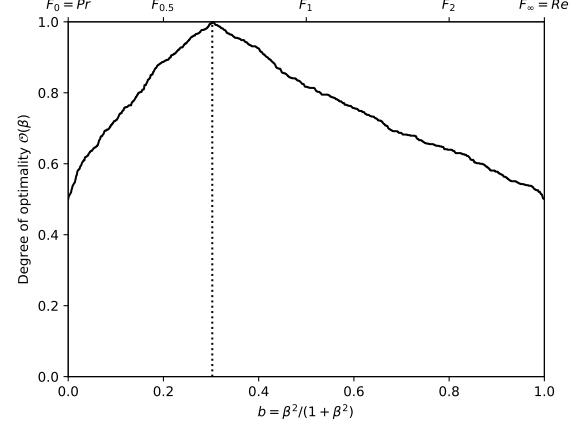
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

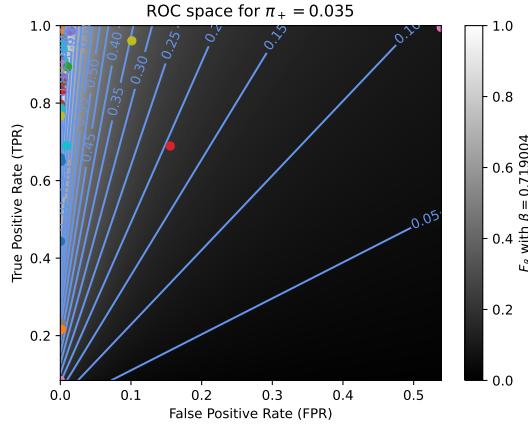


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

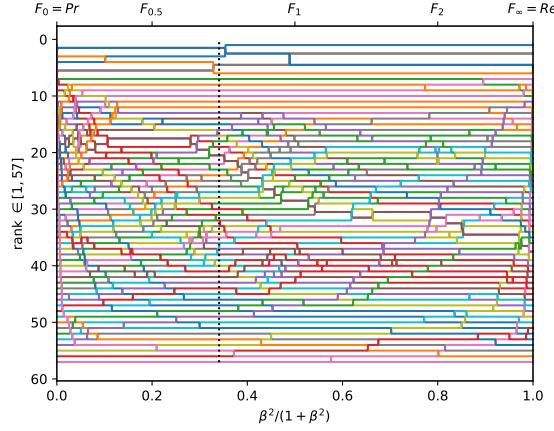


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

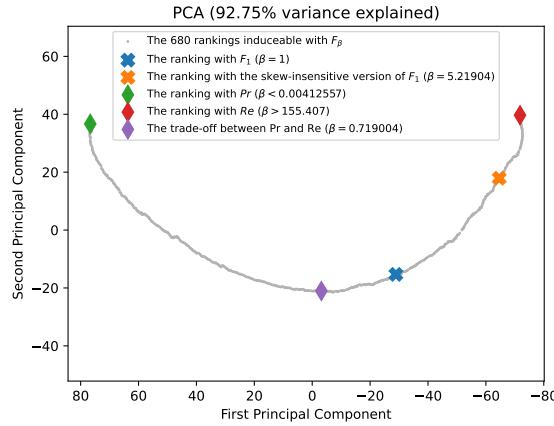
Figure A.3.13. Ranking of 57 BGS methods evaluated on the video "overpass" ($\pi_+ = 0.0134$).



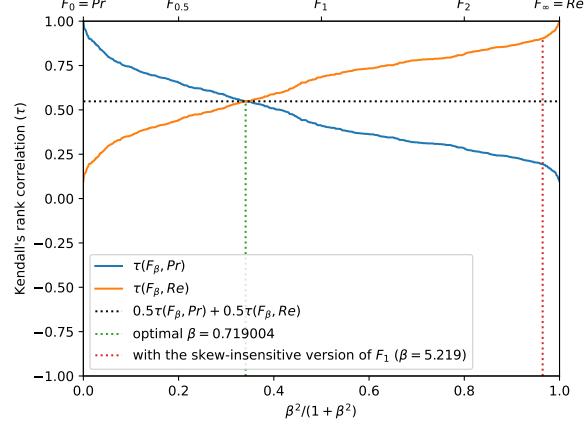
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



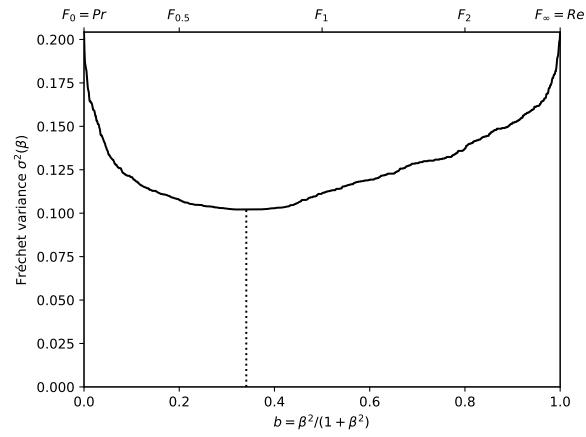
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



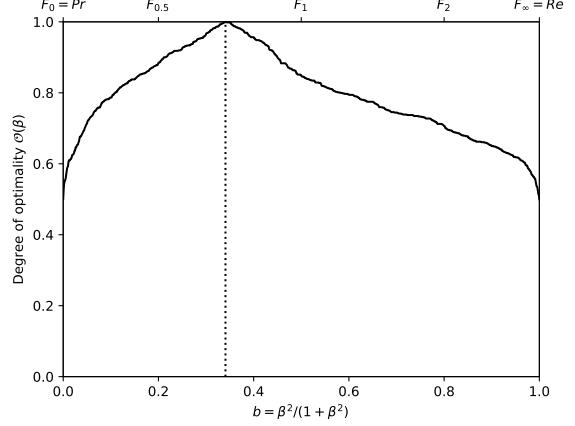
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

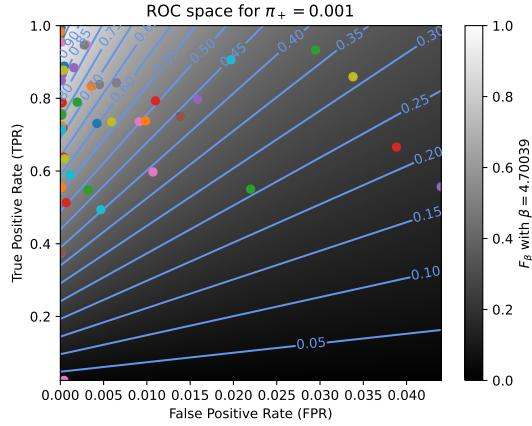


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

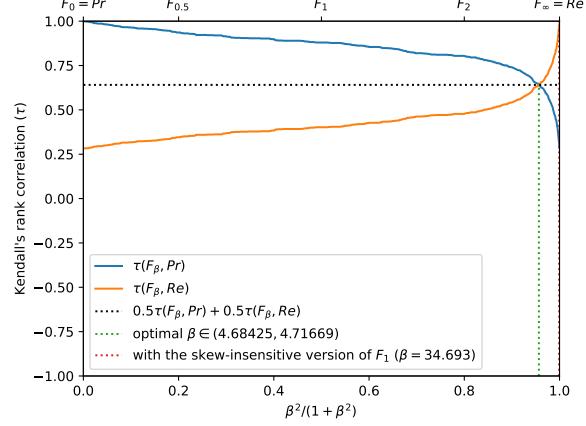


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

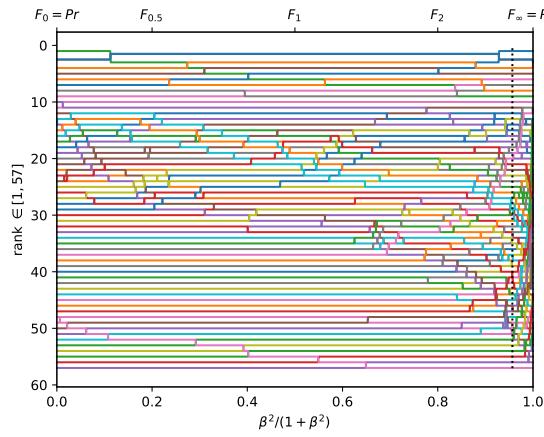
Figure A.3.14. Ranking of 57 BGS methods evaluated on the video "canoe" ($\pi_+ = 0.0354$).



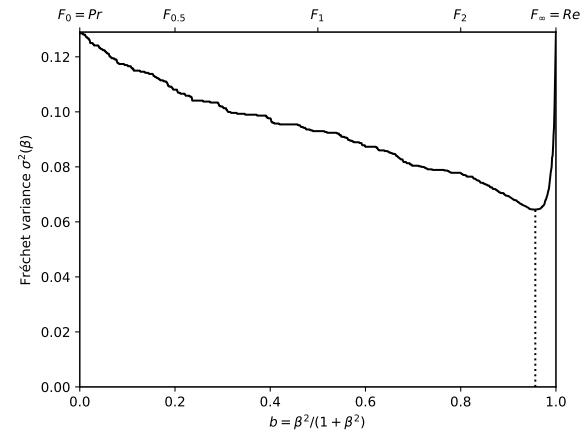
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



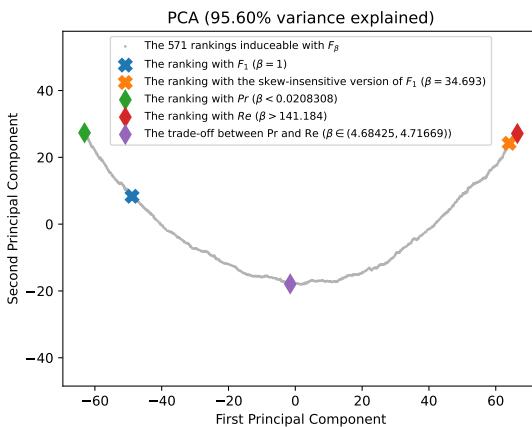
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(Re; F_\beta)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



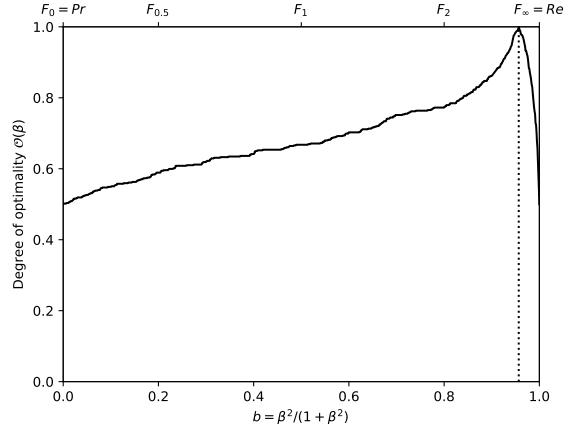
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



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

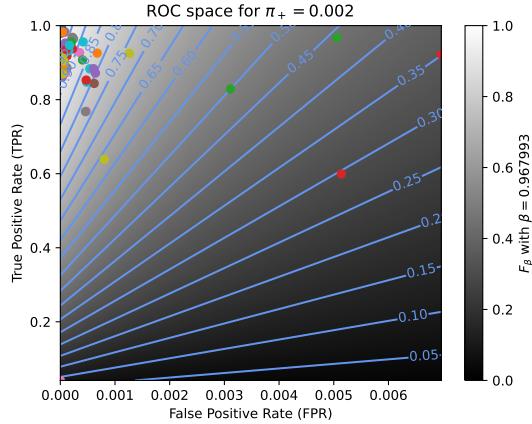


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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

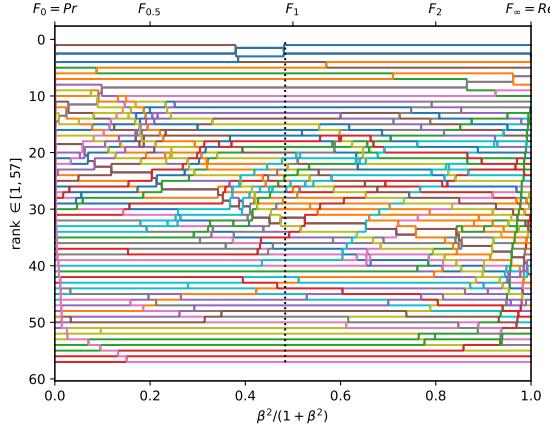


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

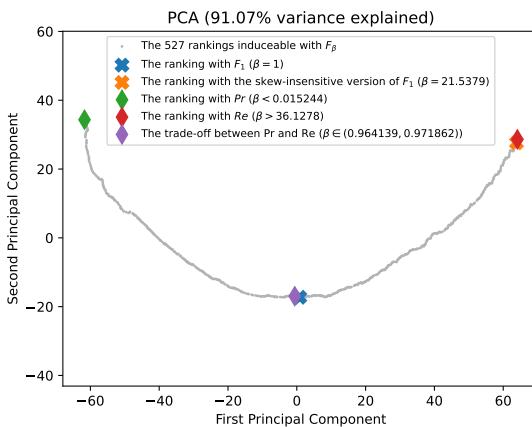
Figure A.3.15. Ranking of 57 BGS methods evaluated on the video "fountain01" ($\pi_+ = 0.0008$).



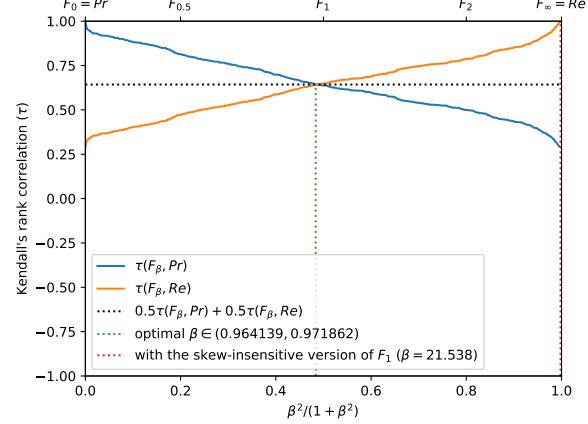
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



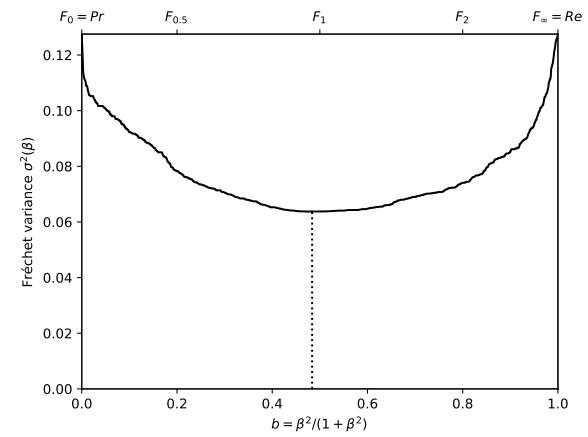
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



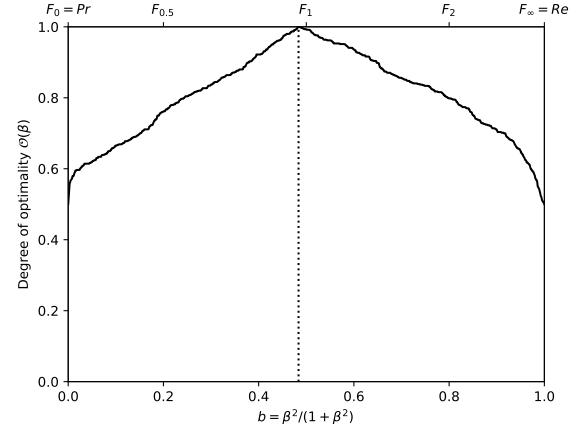
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

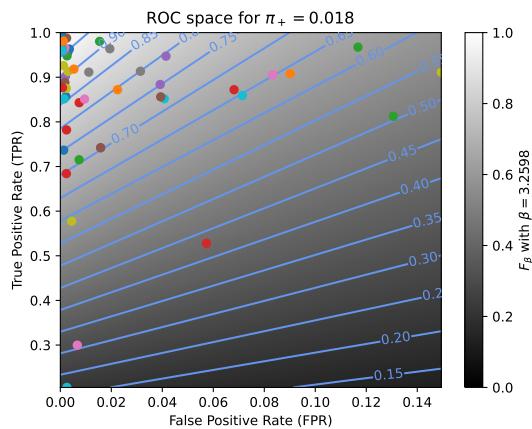


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

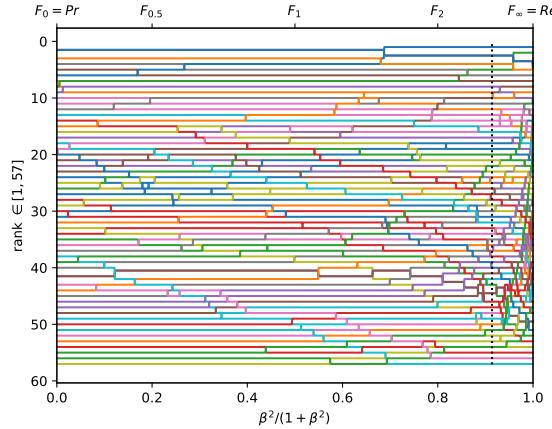


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

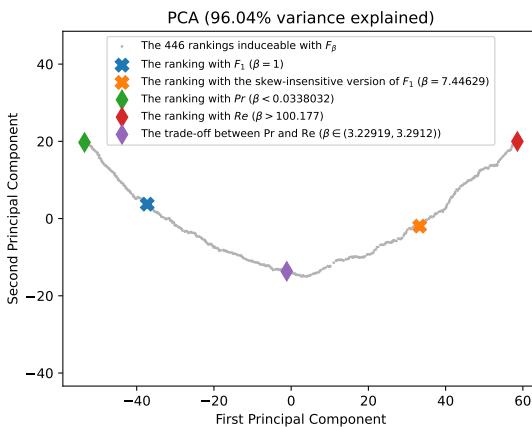
Figure A.3.16. Ranking of 57 BGS methods evaluated on the video "fountain02" ($\pi_+ = 0.0022$).



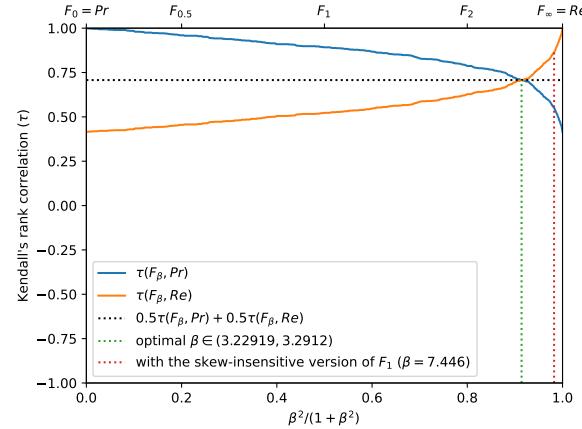
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



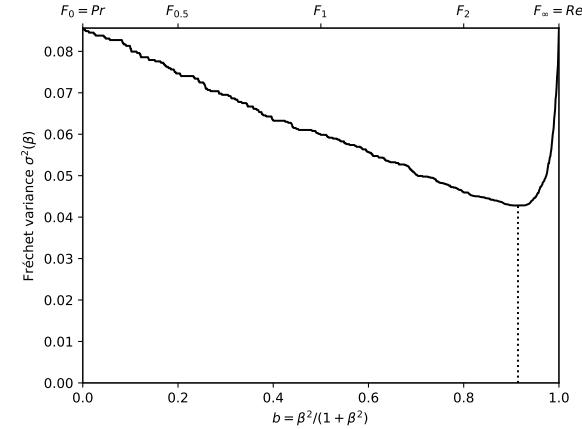
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



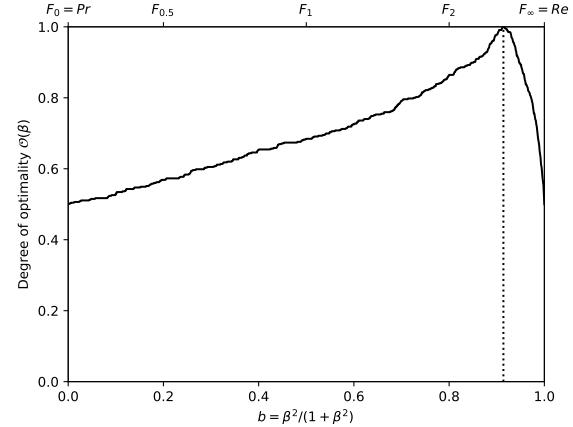
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

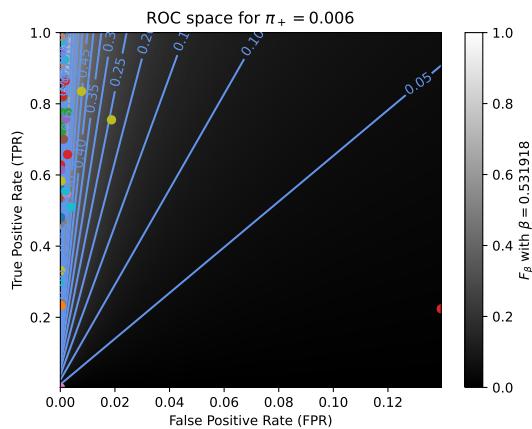


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

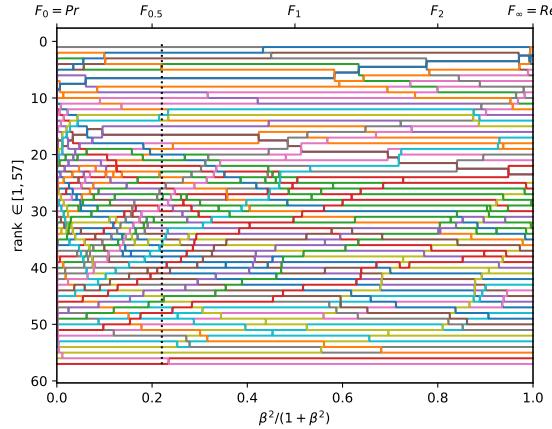


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

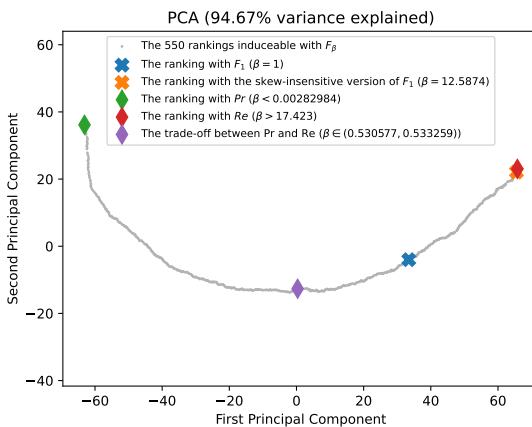
Figure A.3.17. Ranking of 57 BGS methods evaluated on the video "fall" ($\pi_+ = 0.0177$).



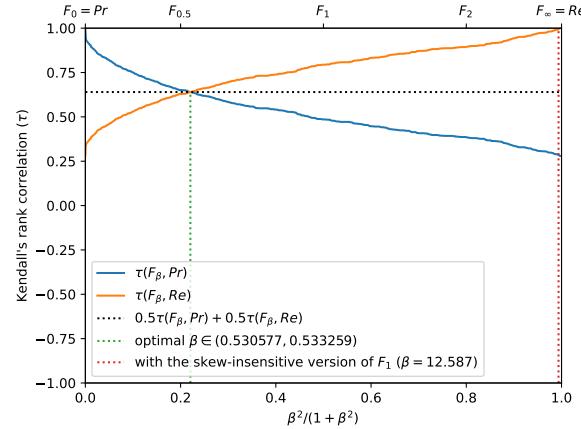
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



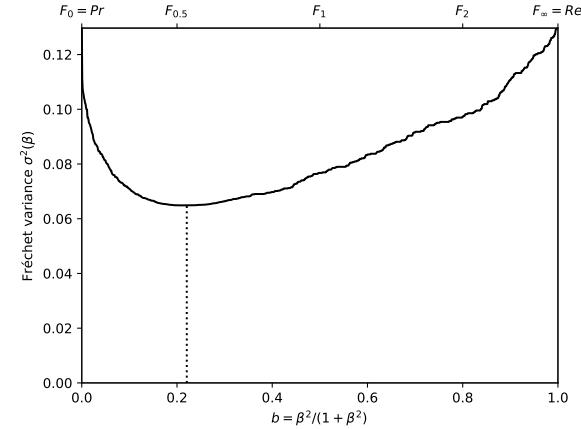
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



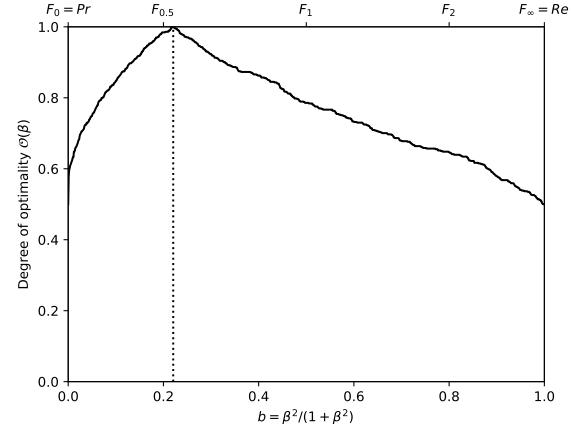
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

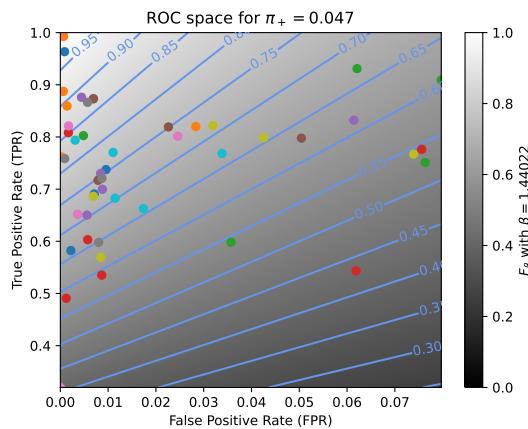


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

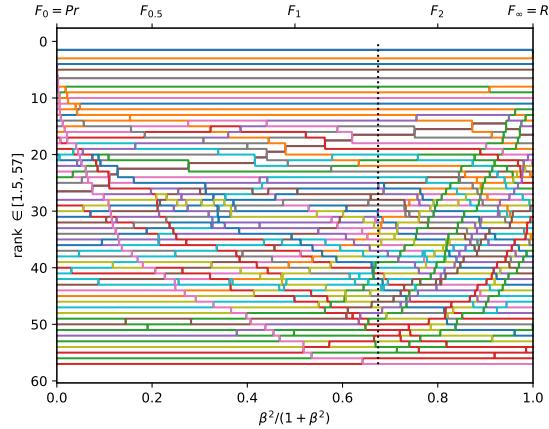


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

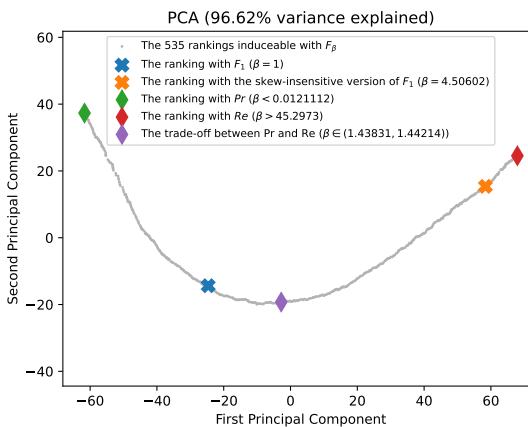
Figure A.3.18. Ranking of 57 BGS methods evaluated on the video "boats" ($\pi_+ = 0.0063$).



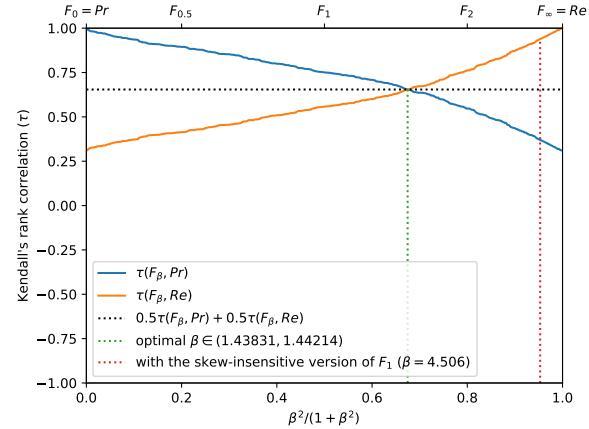
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



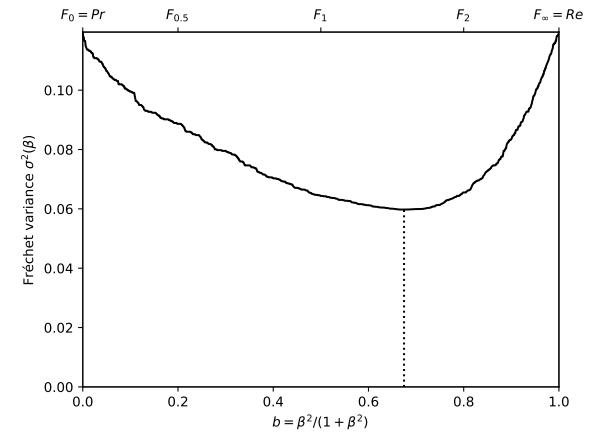
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



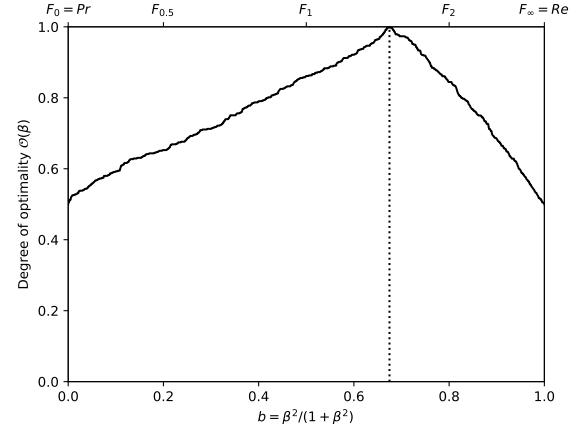
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

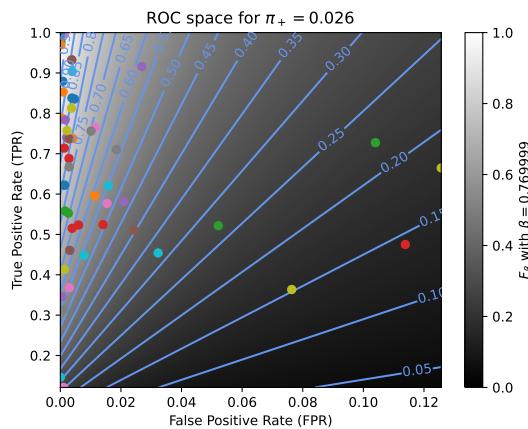


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

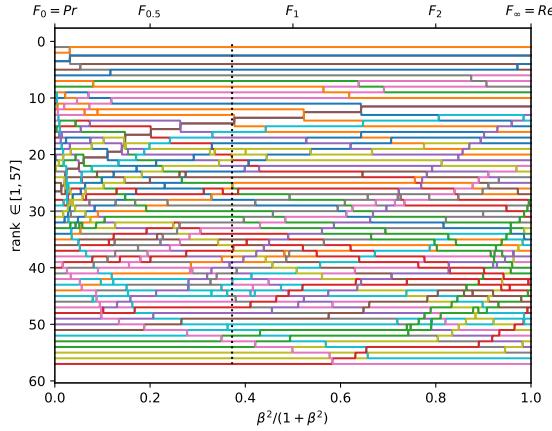


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

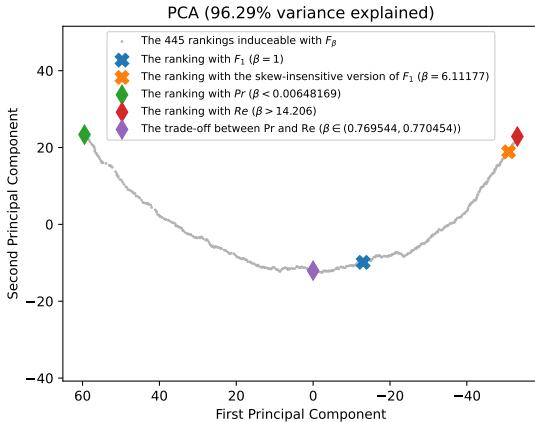
Figure A.3.19. Ranking of 57 BGS methods evaluated on the video "boulevard" ($\pi_+ = 0.0469$).



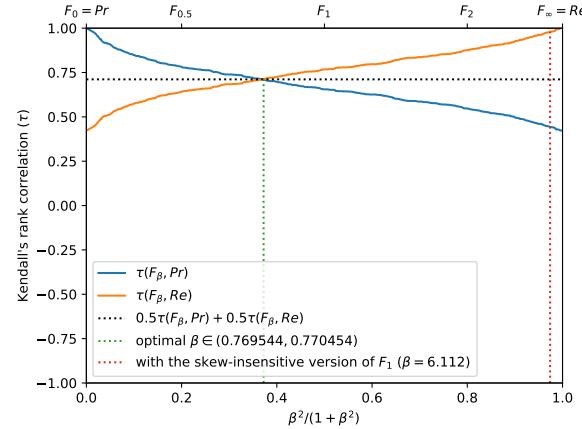
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



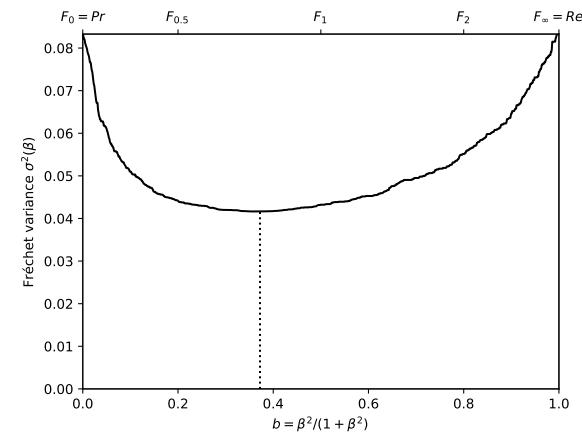
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



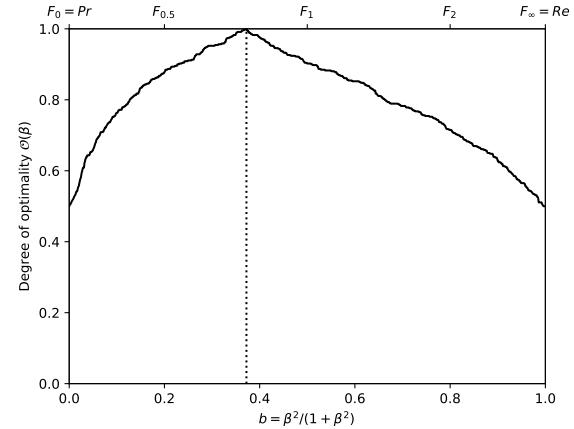
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

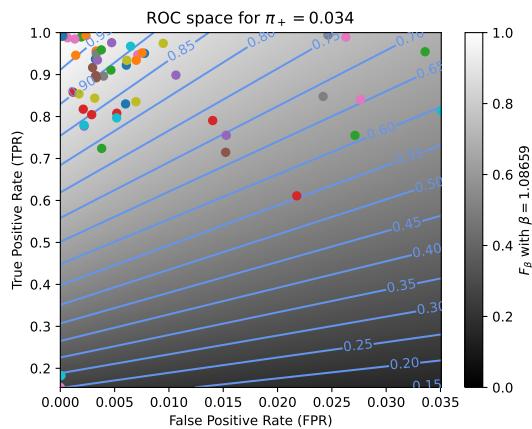


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

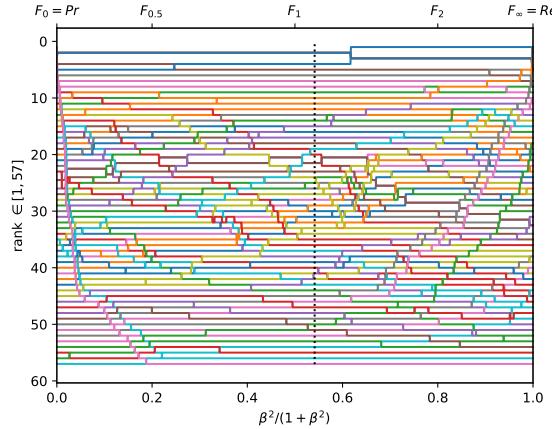


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

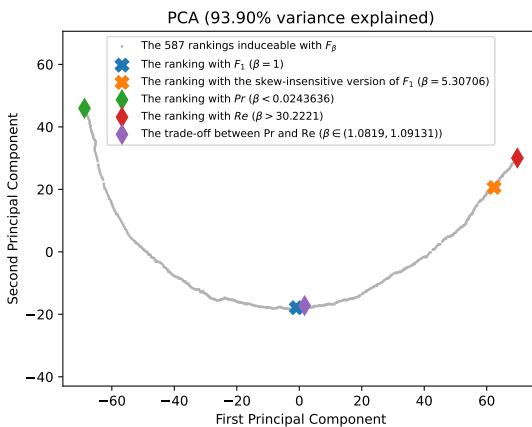
Figure A.3.20. Ranking of 57 BGS methods evaluated on the video "sidewalk" ($\pi_+ = 0.0261$).



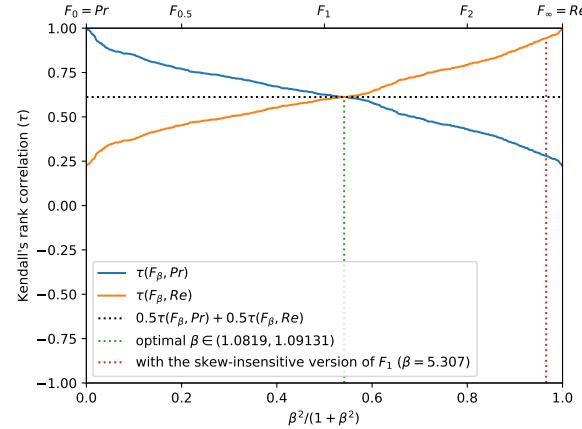
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



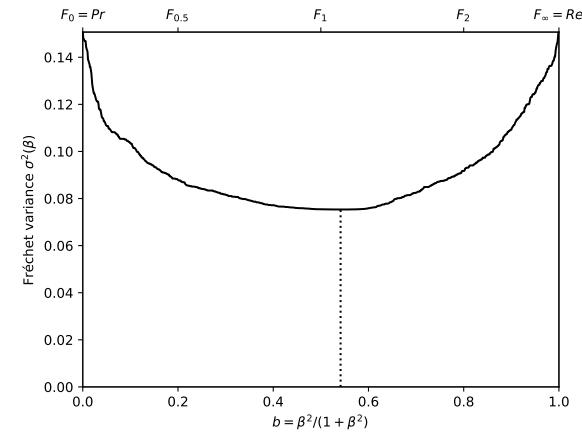
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



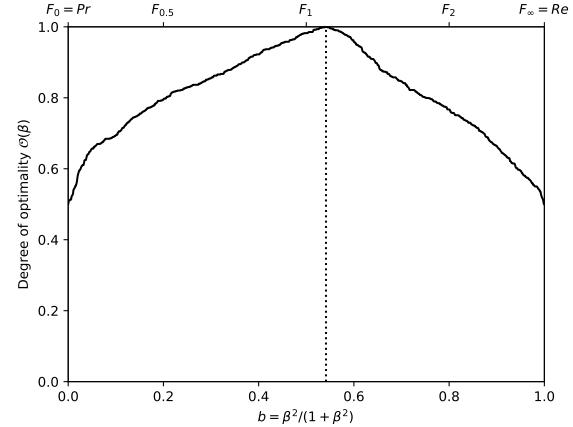
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

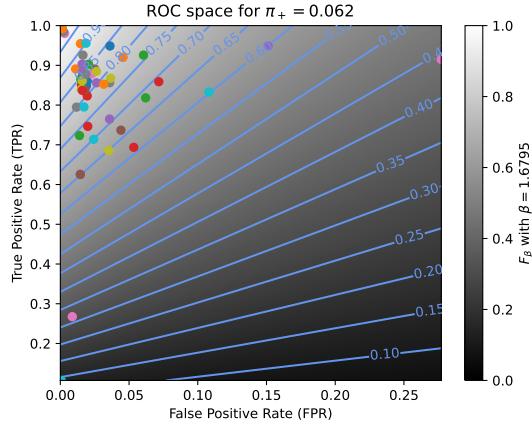


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

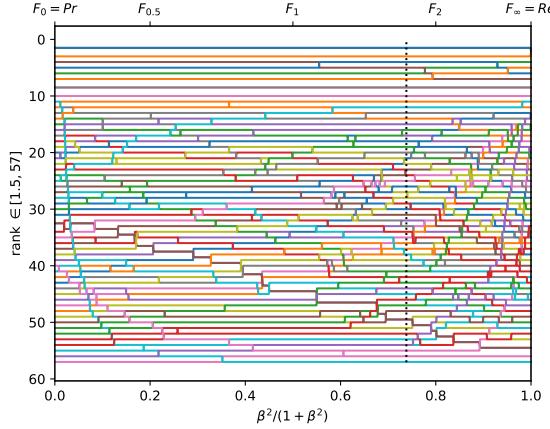


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

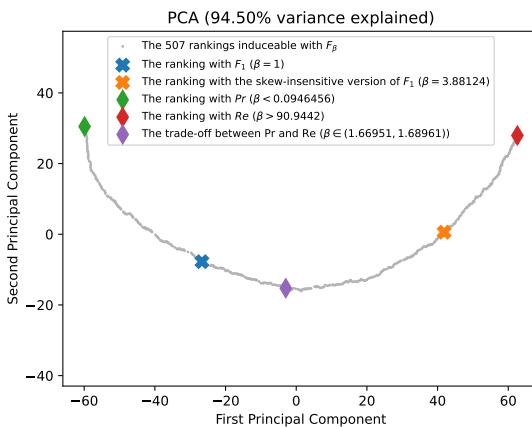
Figure A.3.21. Ranking of 57 BGS methods evaluated on the video "badminton" ($\pi_+ = 0.0343$).



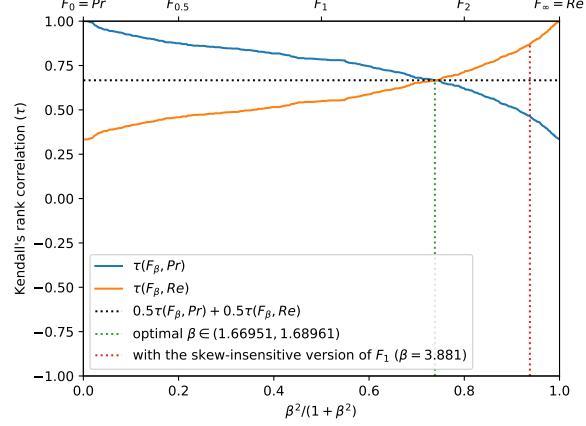
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



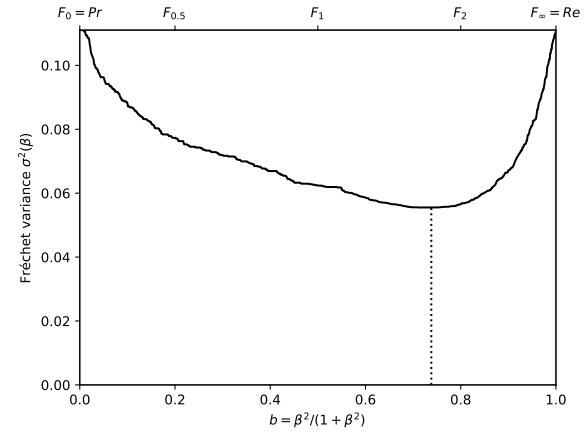
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



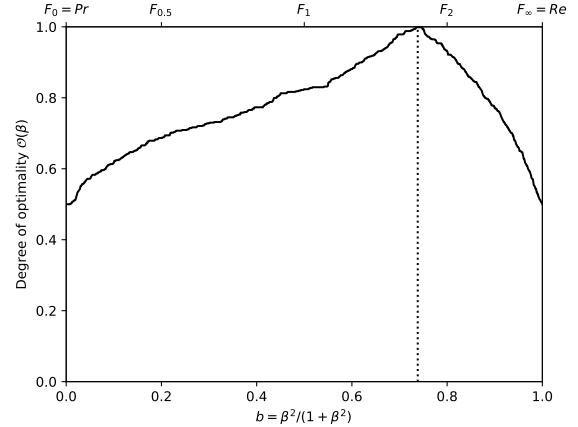
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

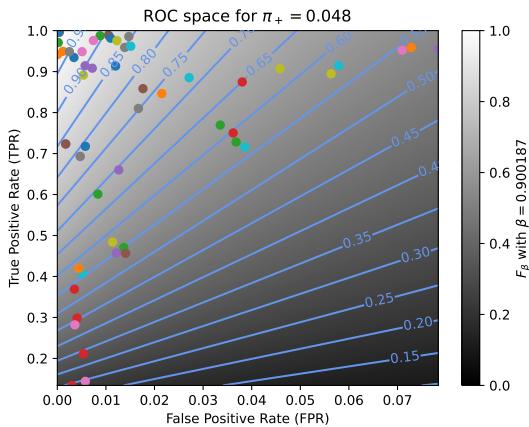


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

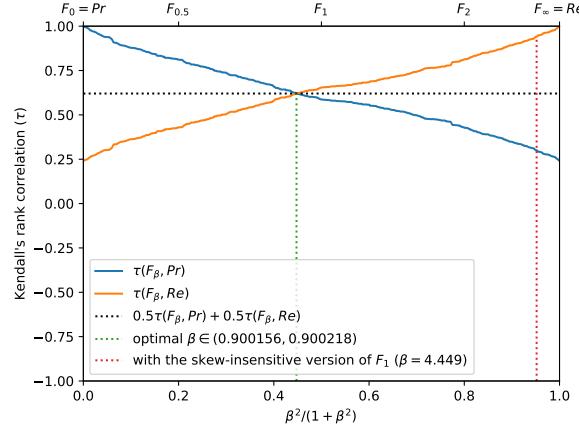


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

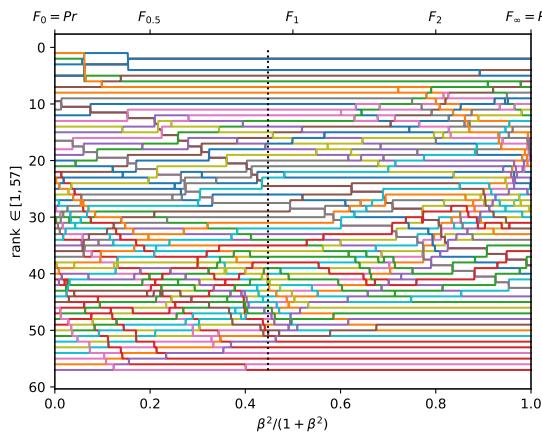
Figure A.3.22. Ranking of 57 BGS methods evaluated on the video "traffic" ($\pi_+ = 0.0623$).



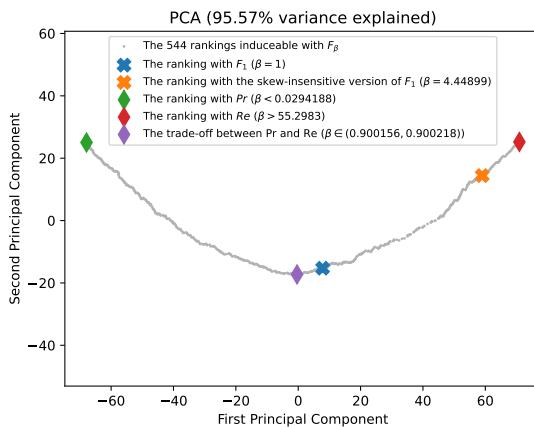
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



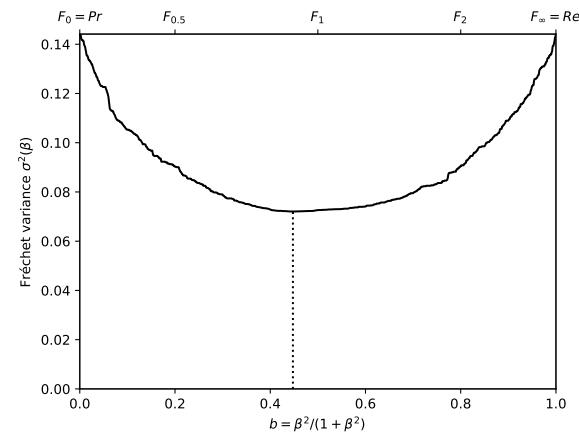
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



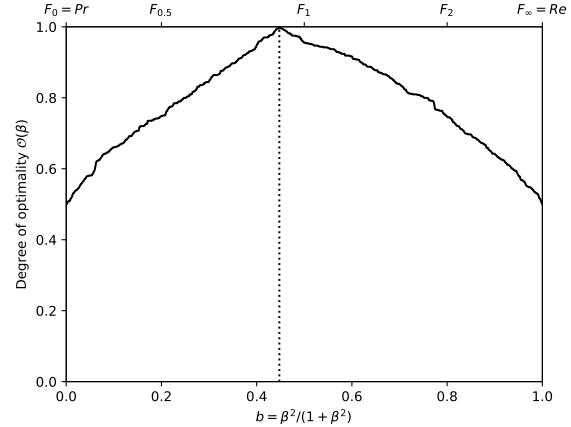
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

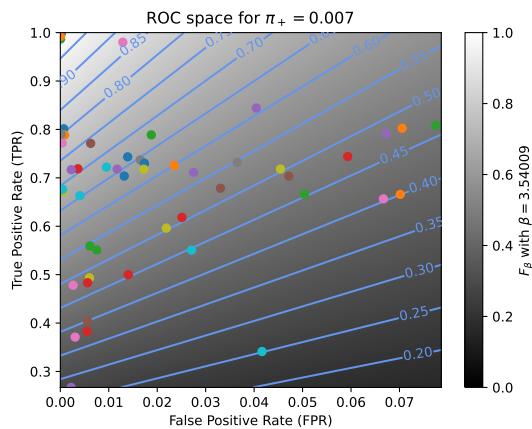


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

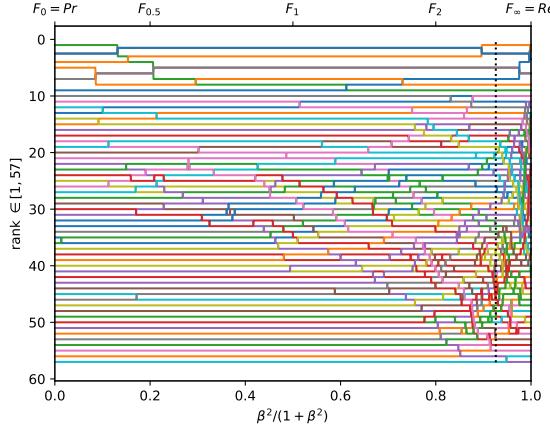


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

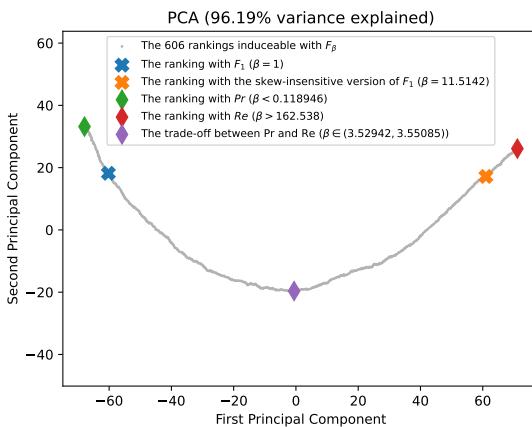
Figure A.3.23. Ranking of 57 BGS methods evaluated on the video "abandonedBox" ($\pi_+ = 0.0481$).



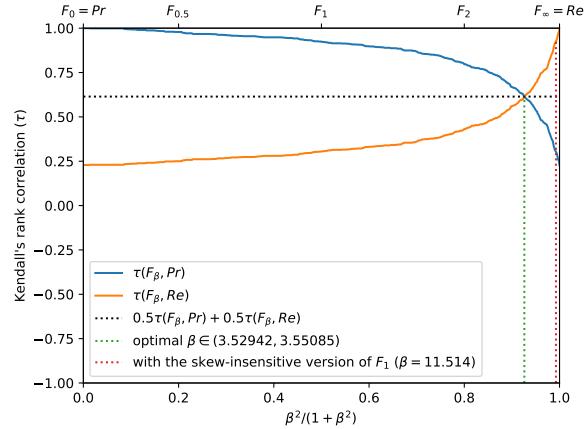
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



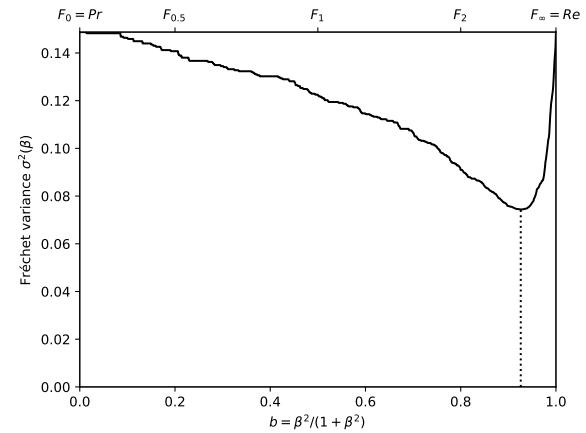
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



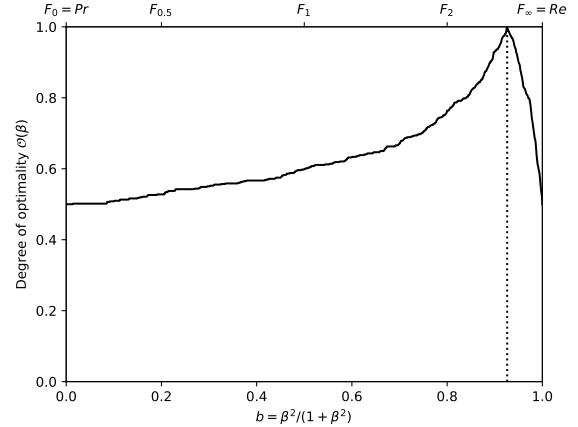
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



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

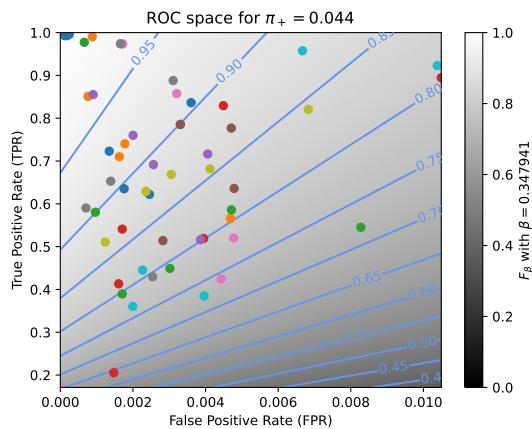


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(Re; F_\beta)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

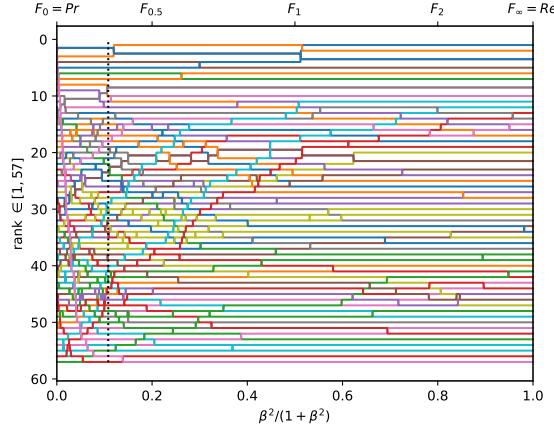


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

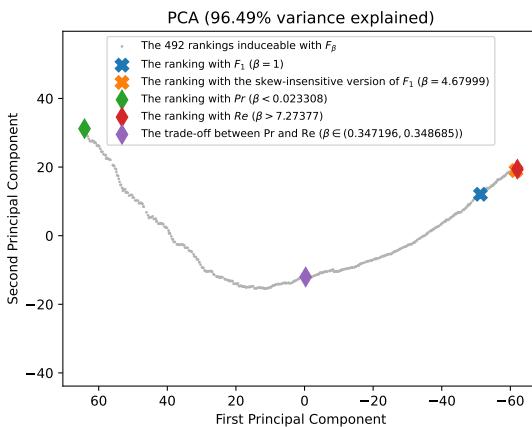
Figure A.3.24. Ranking of 57 BGS methods evaluated on the video "winterDriveway" ($\pi_+ = 0.0075$).



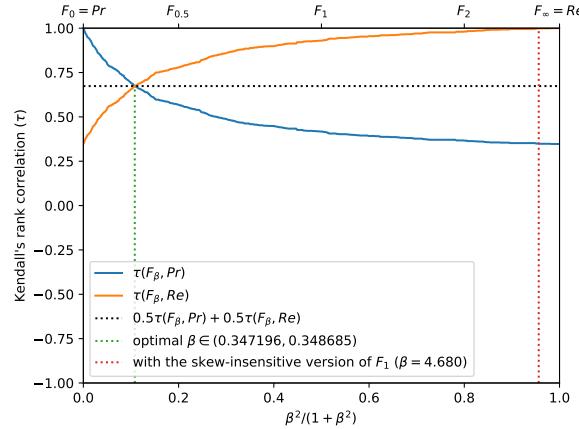
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



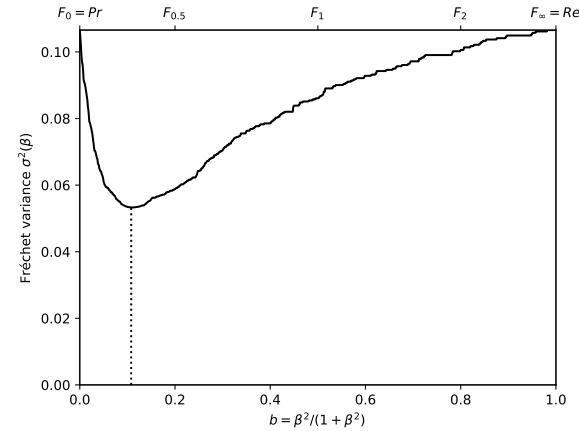
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



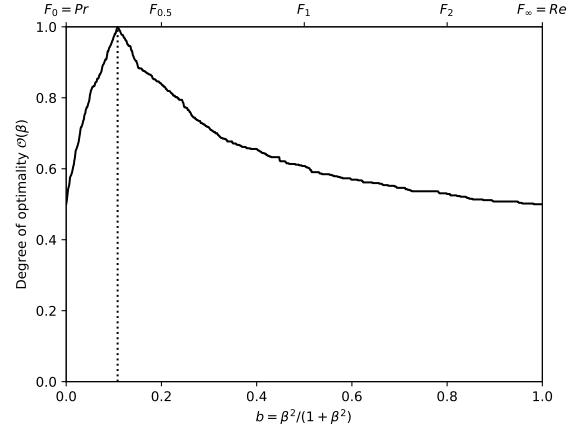
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



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

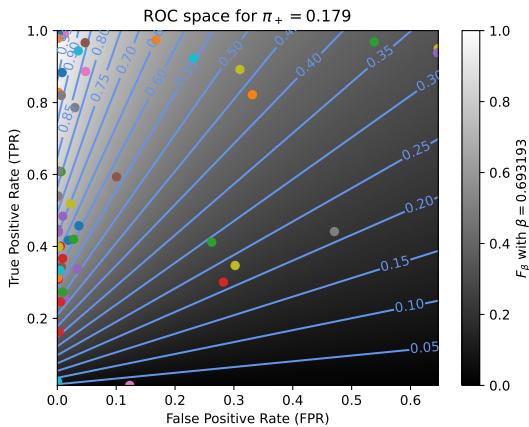


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

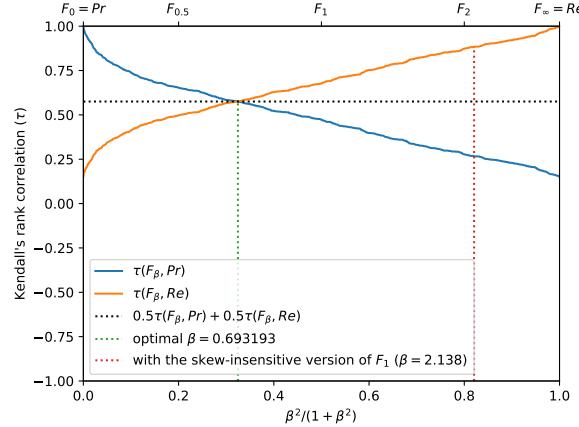


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

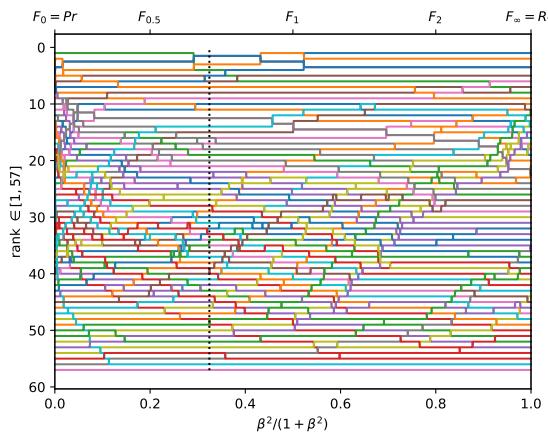
Figure A.3.25. Ranking of 57 BGS methods evaluated on the video "sofa" ($\pi_+ = 0.0437$).



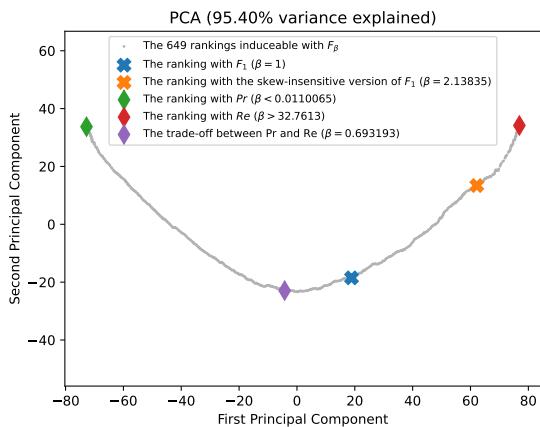
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



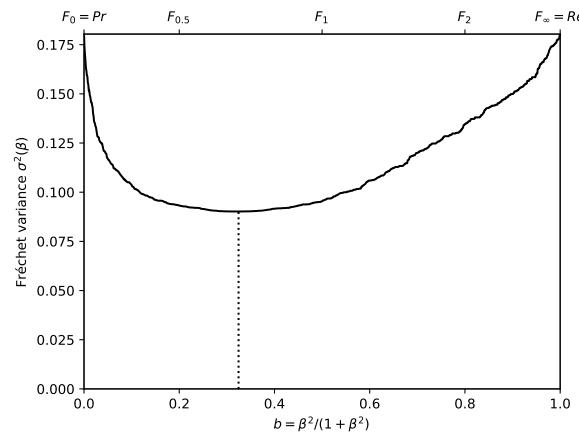
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



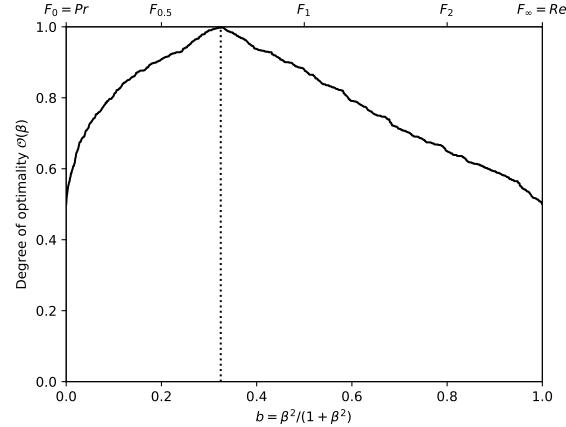
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

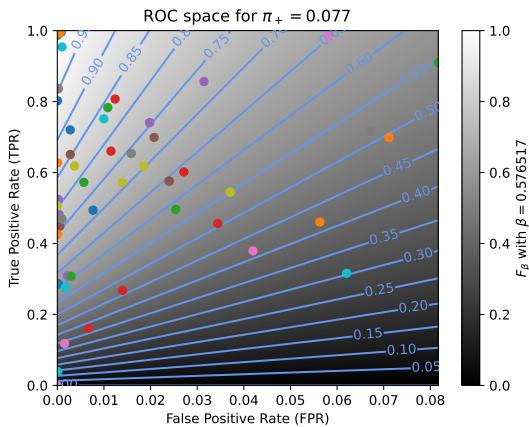


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

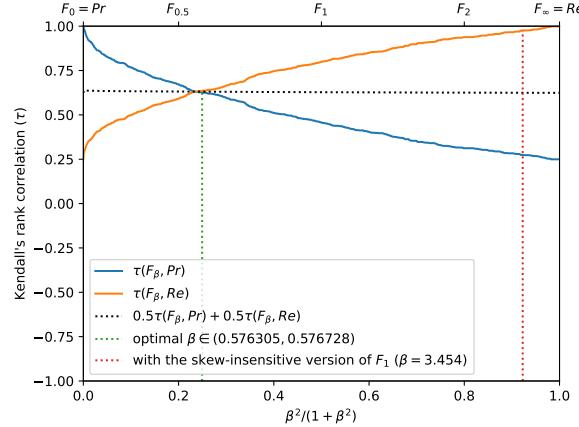


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

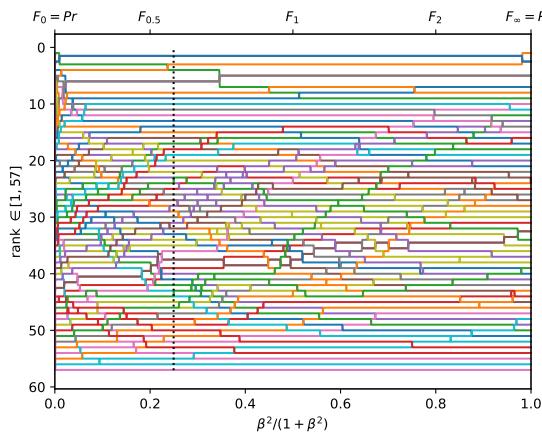
Figure A.3.26. Ranking of 57 BGS methods evaluated on the video "tramstop" ($\pi_+ = 0.1795$).



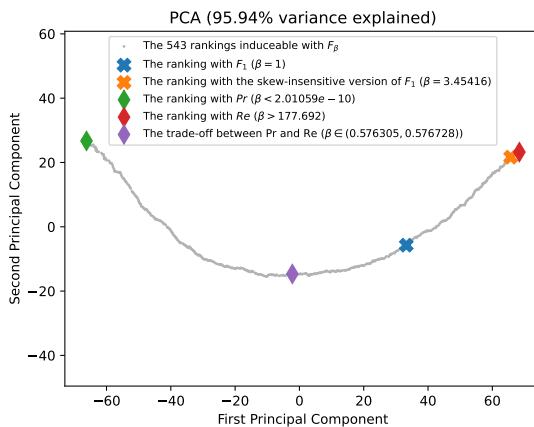
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



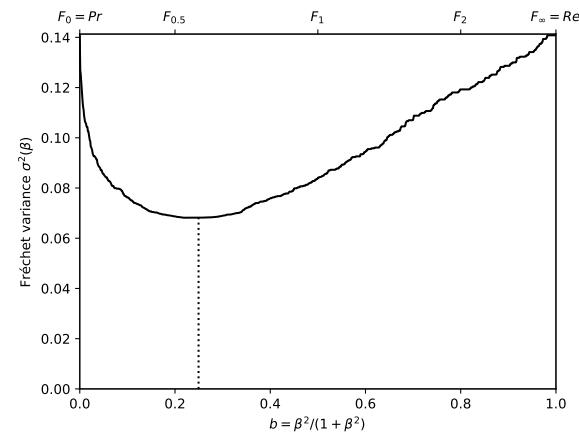
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



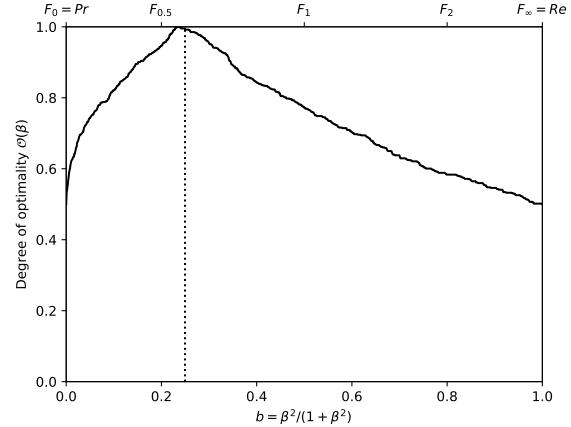
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

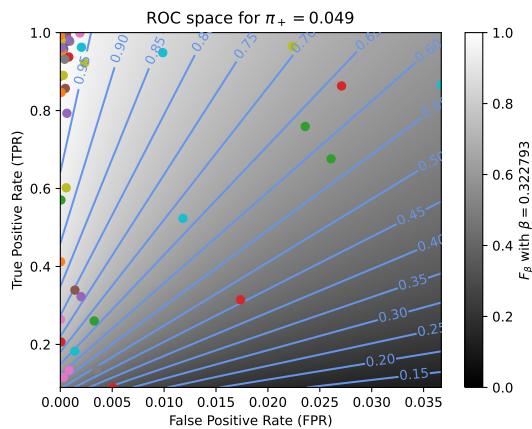


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

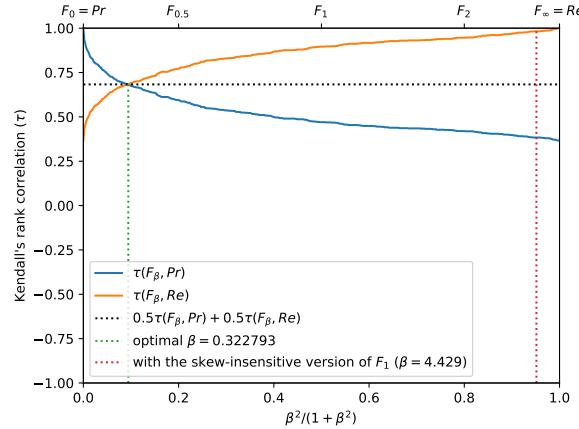


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

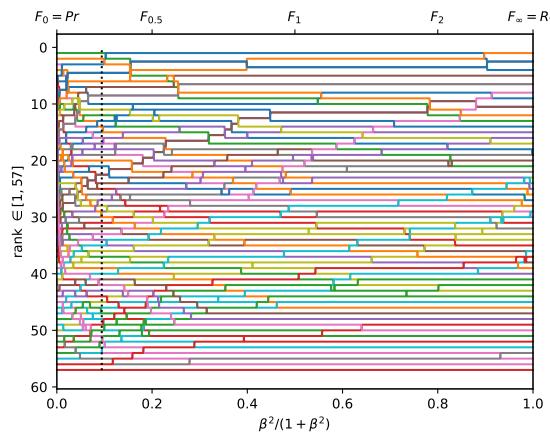
Figure A.3.27. Ranking of 57 BGS methods evaluated on the video "parking" ($\pi_+ = 0.0773$).



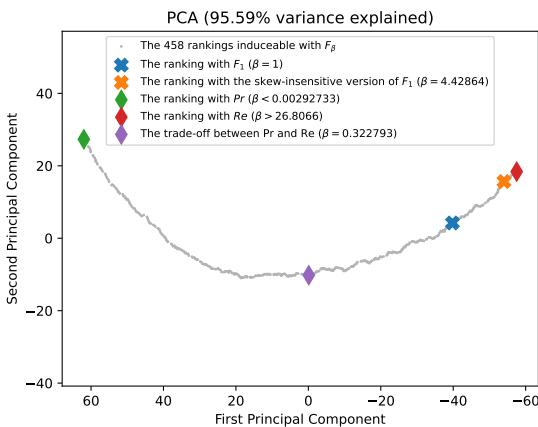
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



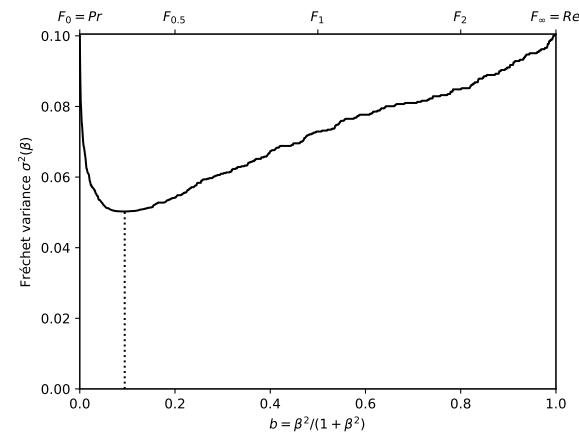
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



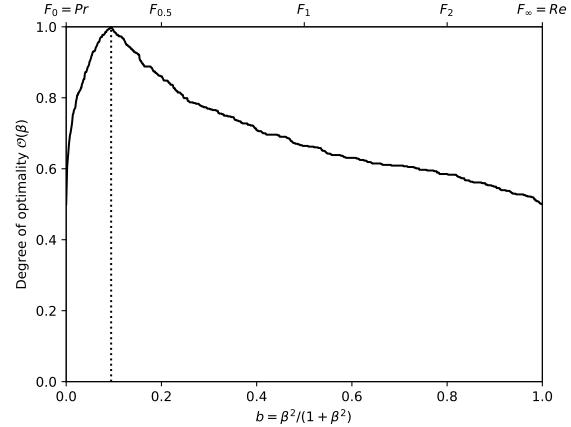
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

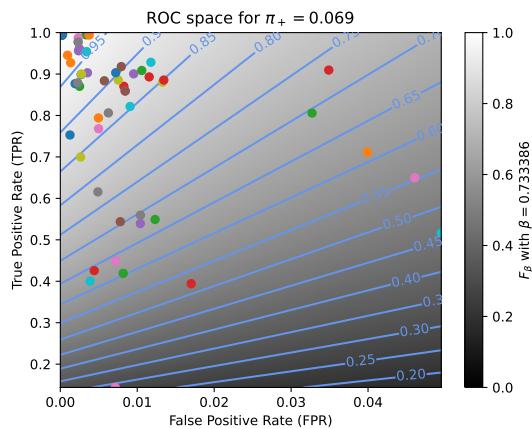


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

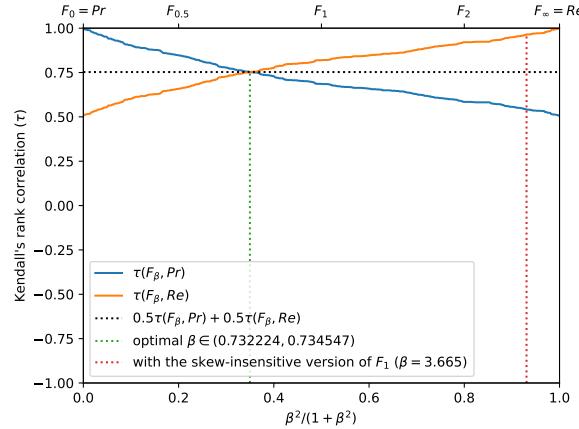


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

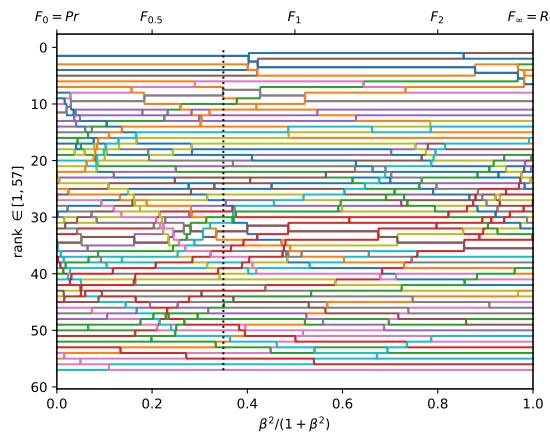
Figure A.3.28. Ranking of 57 BGS methods evaluated on the video "streetLight" ($\pi_+ = 0.0485$).



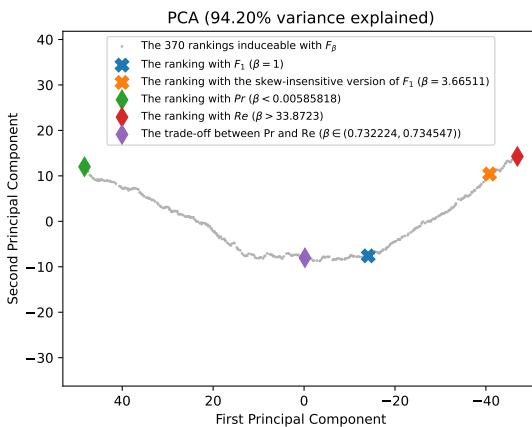
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



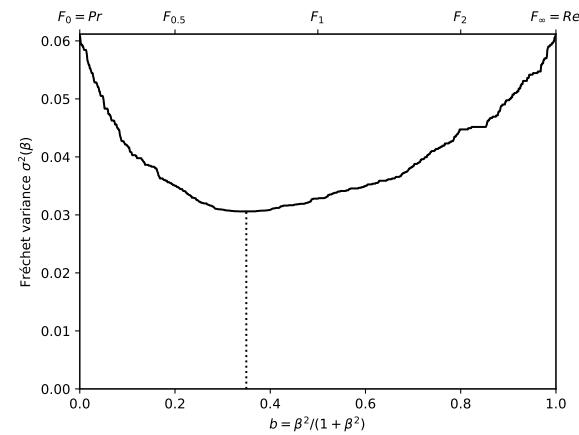
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



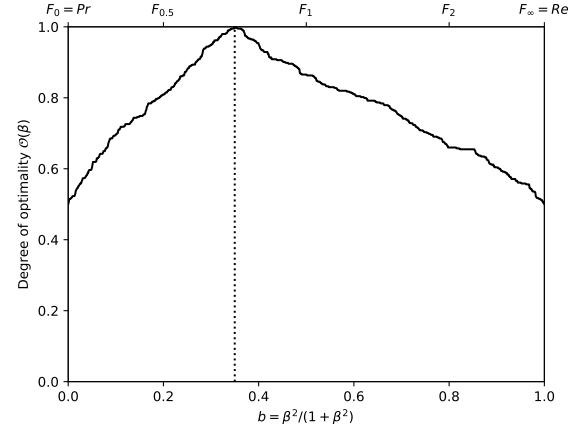
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

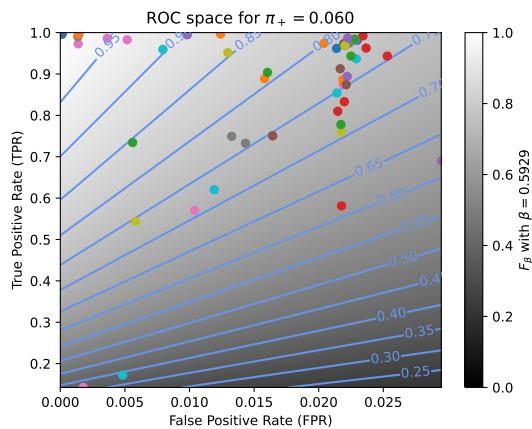


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

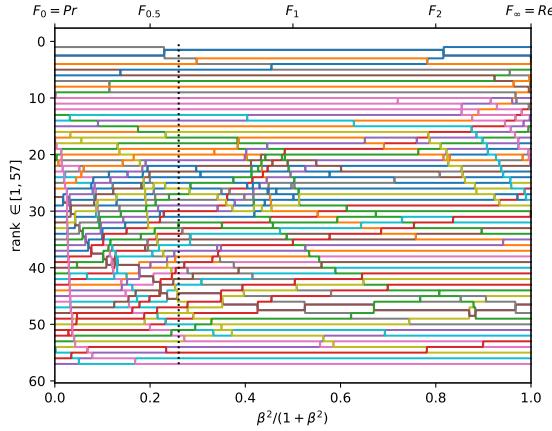


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

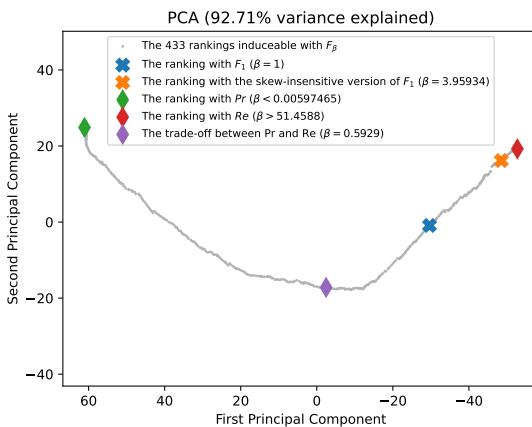
Figure A.3.29. Ranking of 57 BGS methods evaluated on the video "copyMachine" ($\pi_+ = 0.0693$).



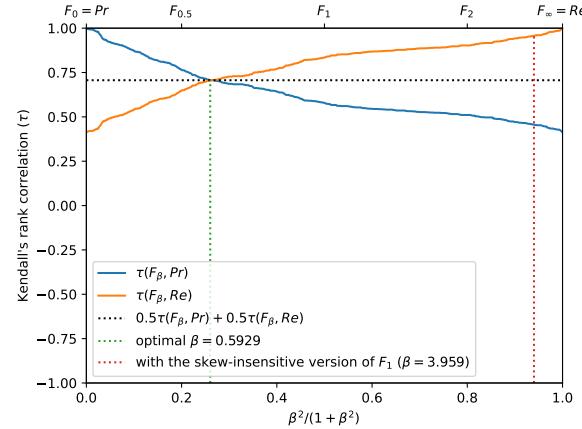
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



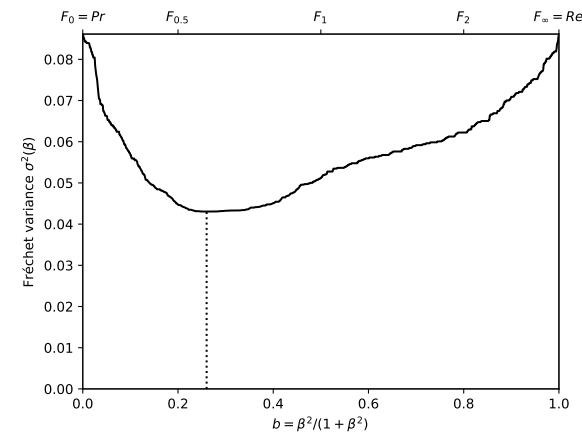
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



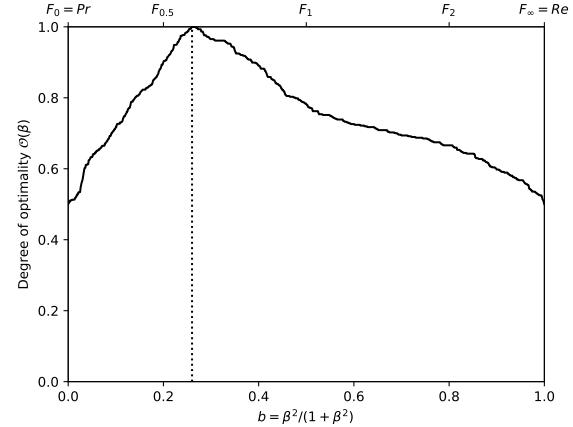
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

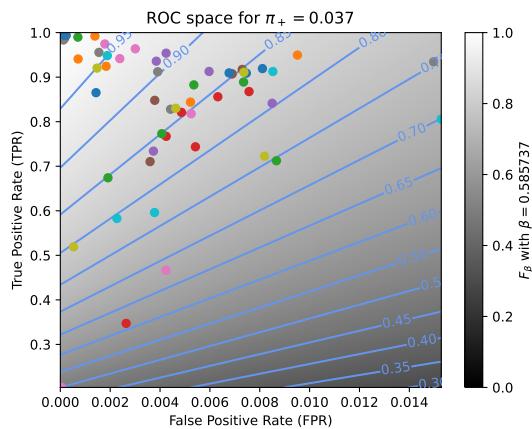


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

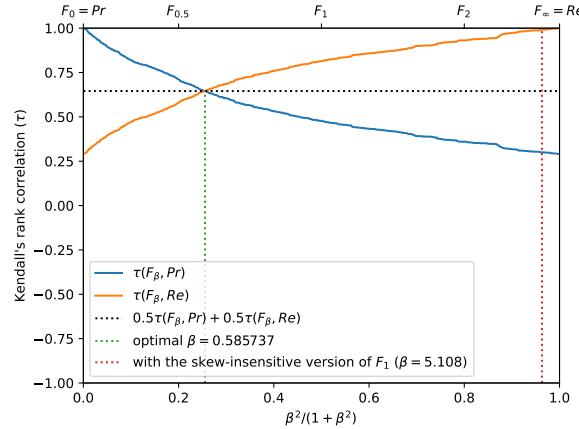


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

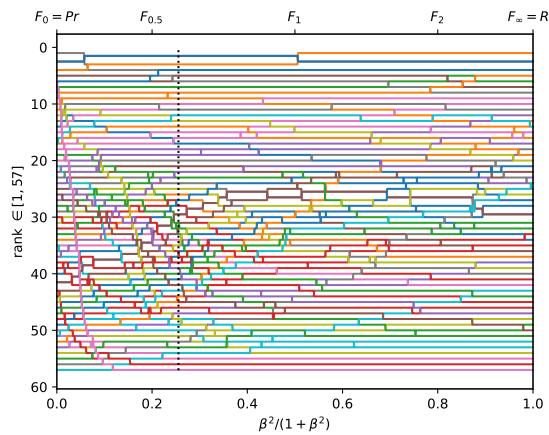
Figure A.3.30. Ranking of 57 BGS methods evaluated on the video "bungalows" ($\pi_+ = 0.0600$).



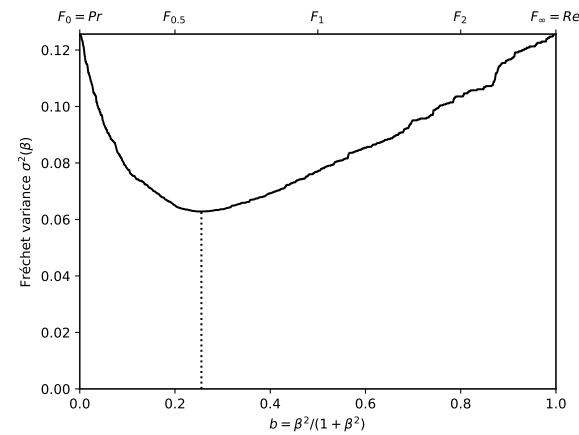
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



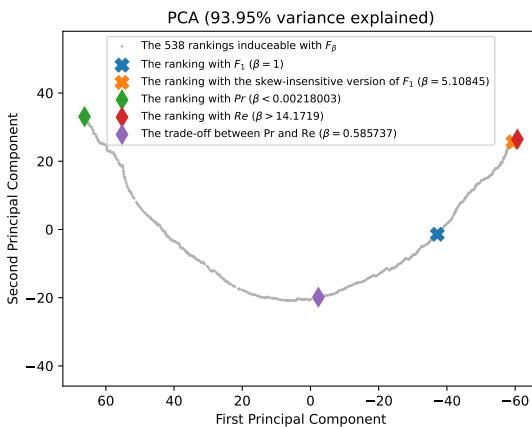
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



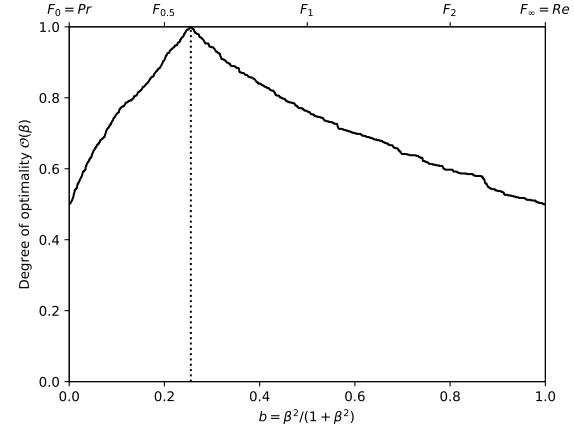
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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.

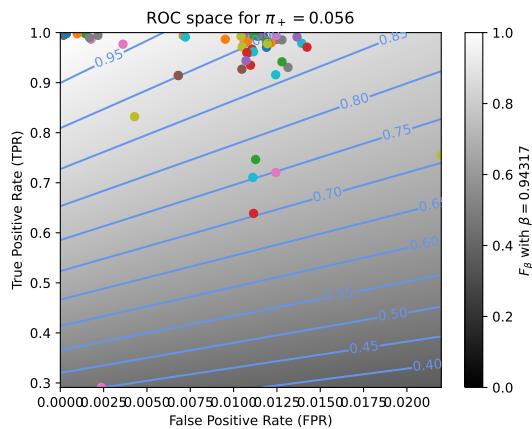


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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

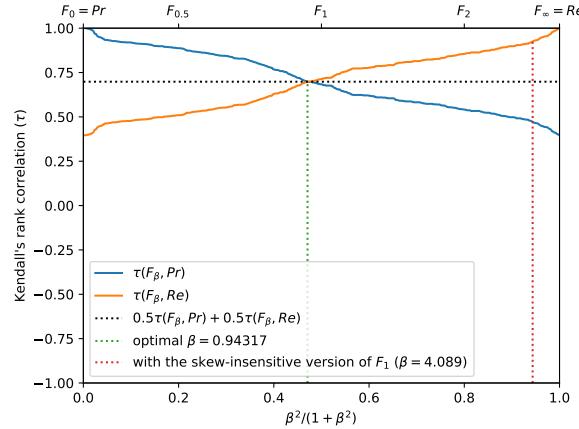


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

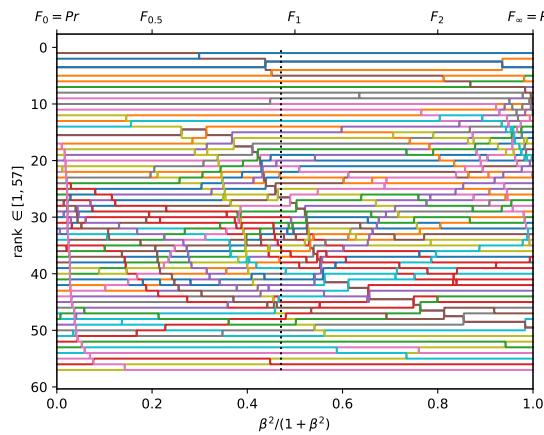
Figure A.3.31. Ranking of 57 BGS methods evaluated on the video "busStation" ($\pi_+ = 0.0369$).



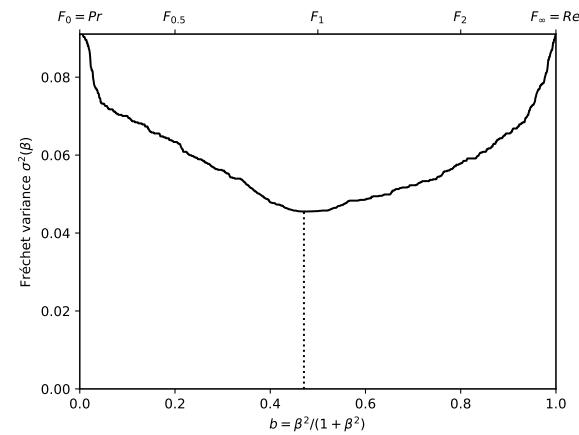
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



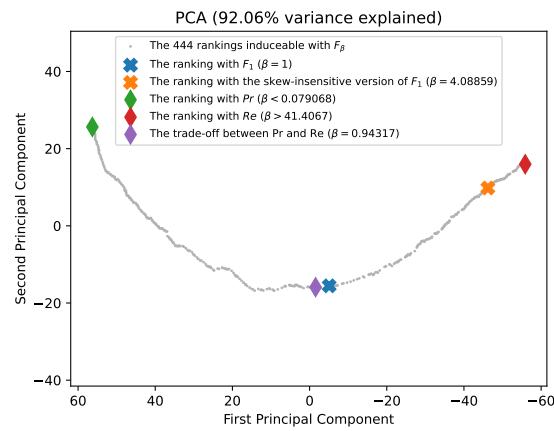
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



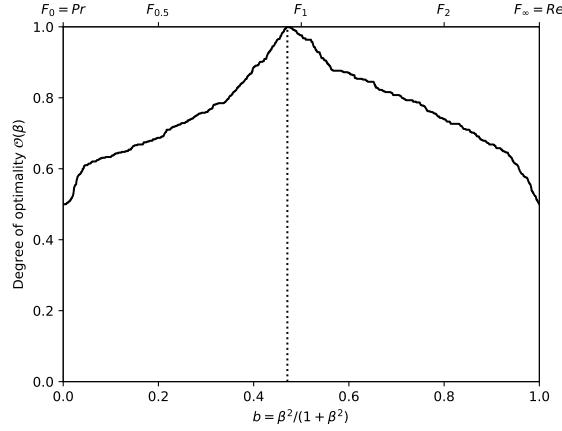
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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.

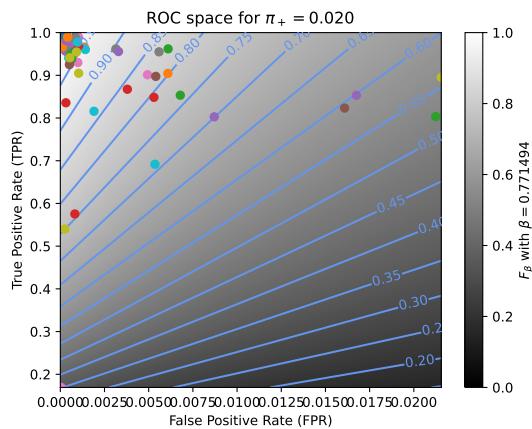


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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

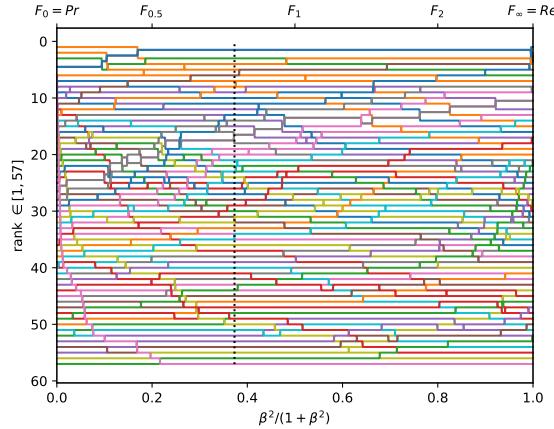


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

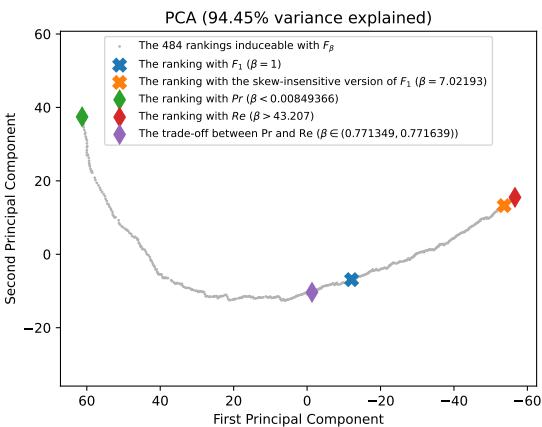
Figure A.3.32. Ranking of 57 BGS methods evaluated on the video "peopleInShade" ($\pi_+ = 0.0564$).



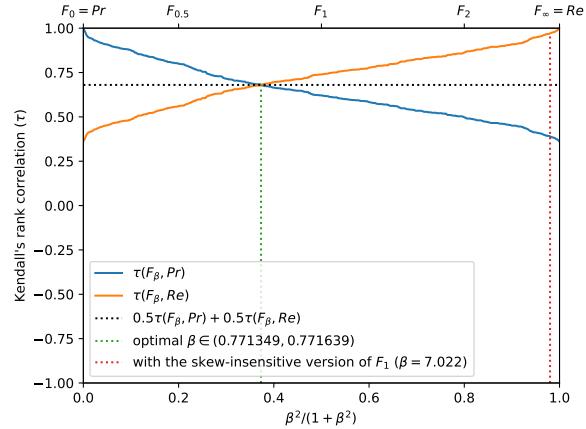
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



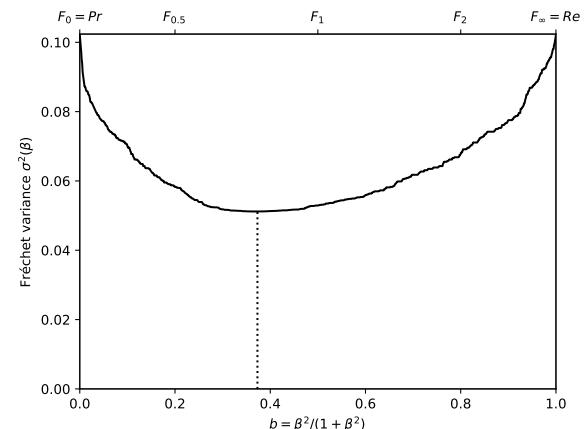
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



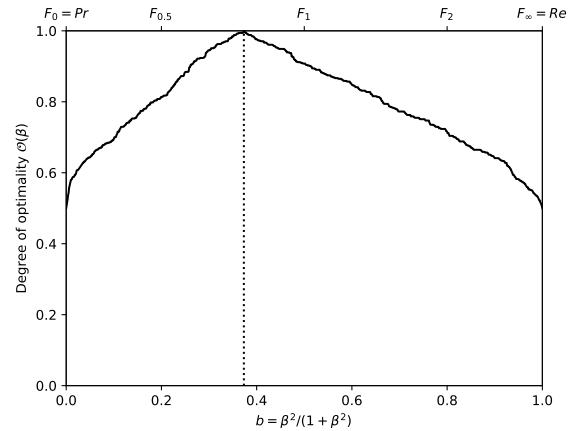
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

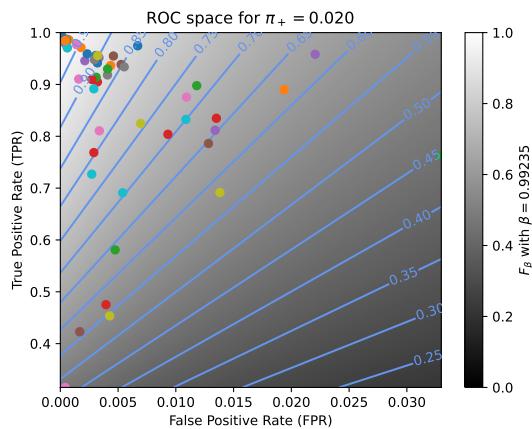


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

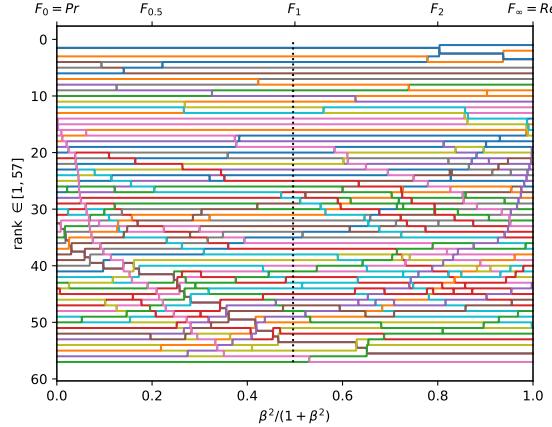


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

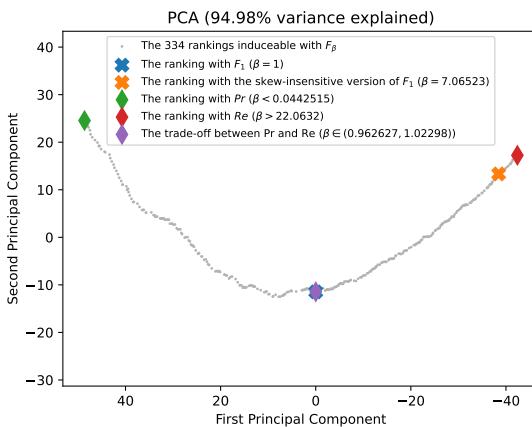
Figure A.3.33. Ranking of 57 BGS methods evaluated on the video "backdoor" ($\pi_+ = 0.0199$).



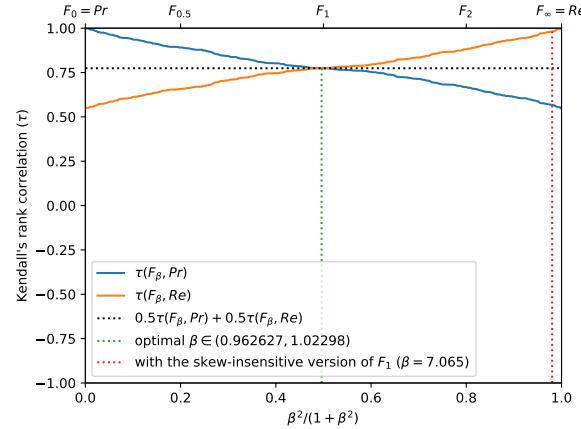
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



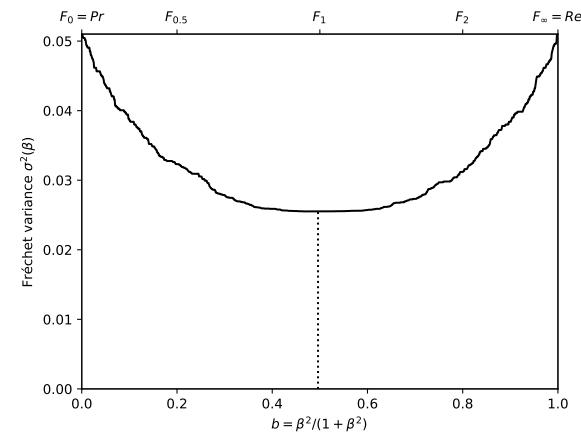
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



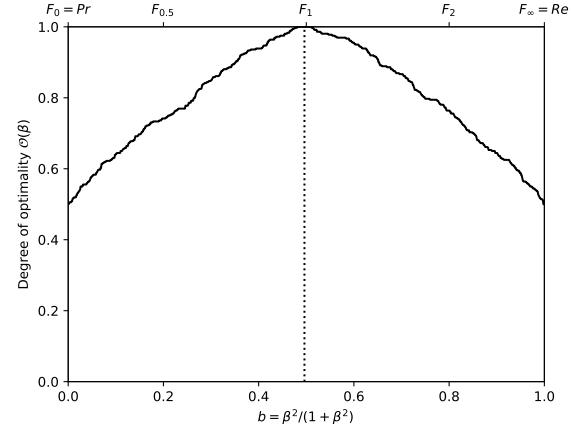
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

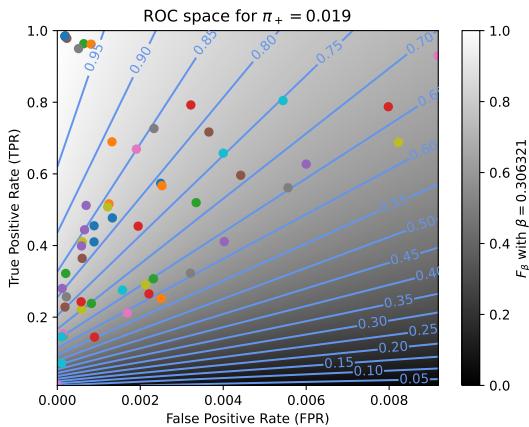


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

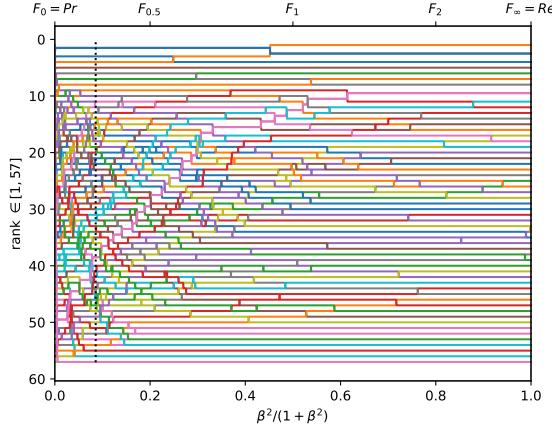


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

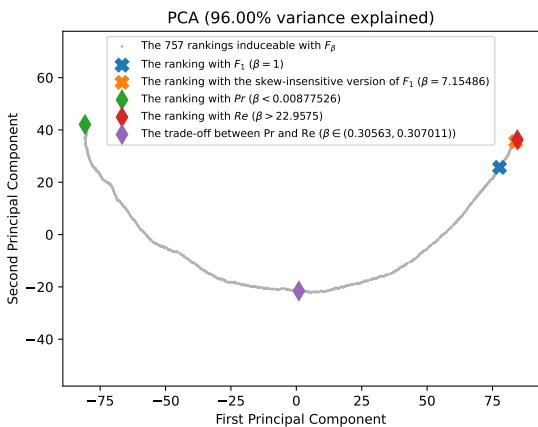
Figure A.3.34. Ranking of 57 BGS methods evaluated on the video "cubicle" ($\pi_+ = 0.0196$).



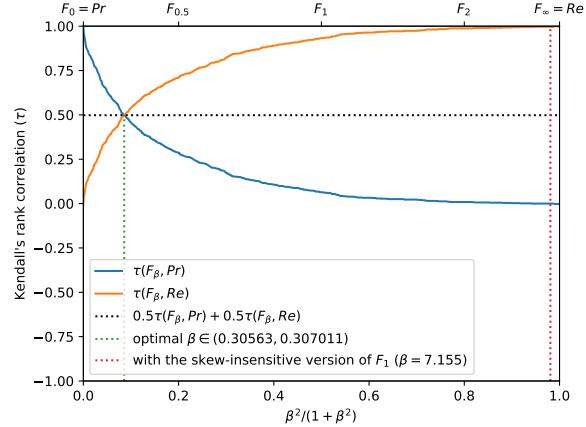
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



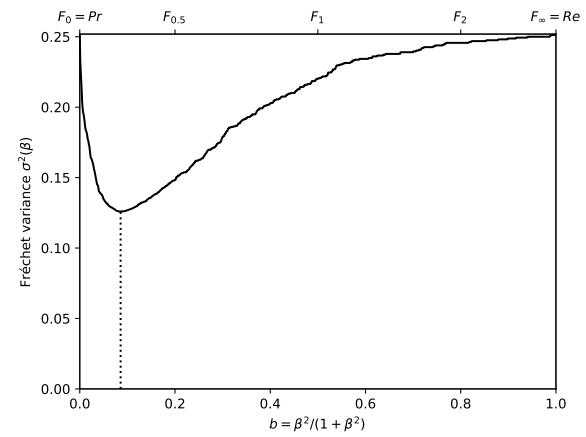
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



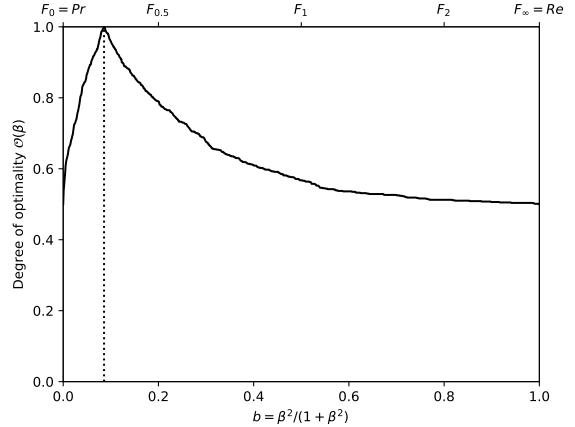
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

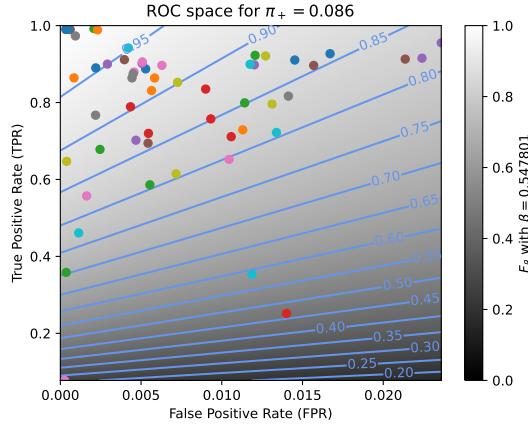


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

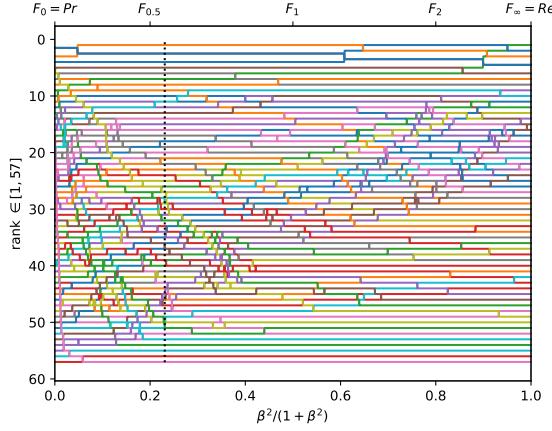


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

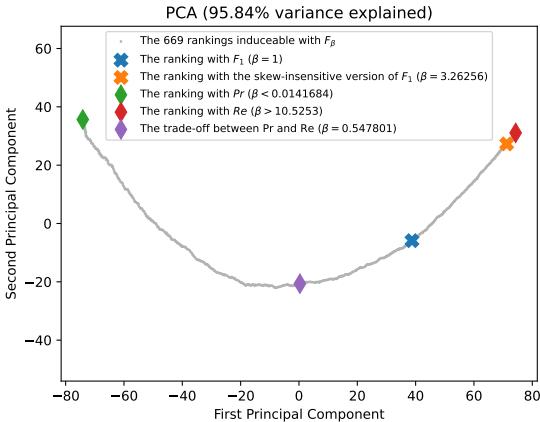
Figure A.3.35. Ranking of 57 BGS methods evaluated on the video "lakeSide" ($\pi_+ = 0.0192$).



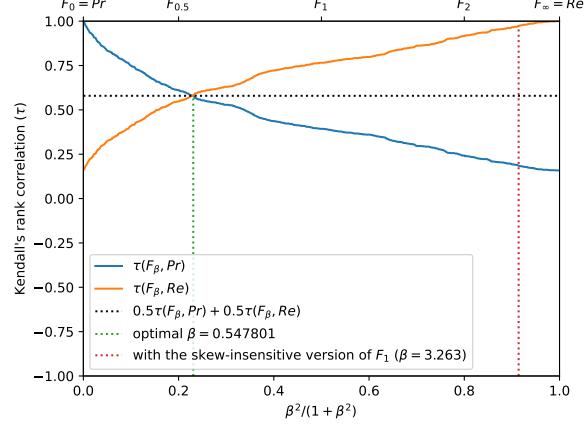
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



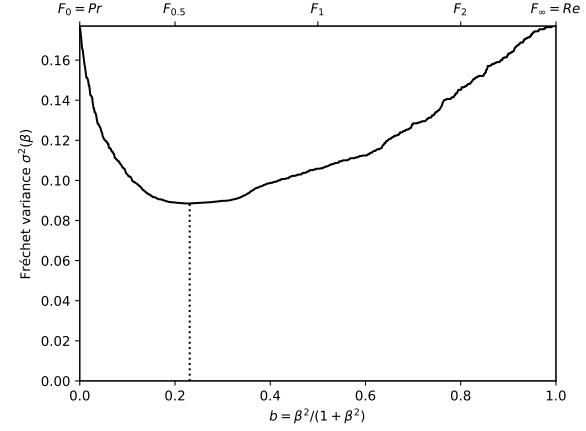
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



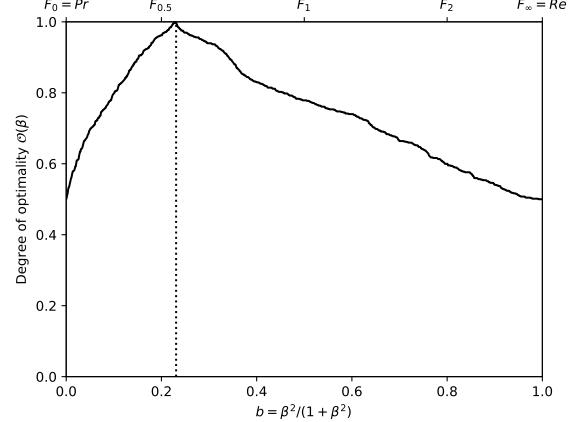
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

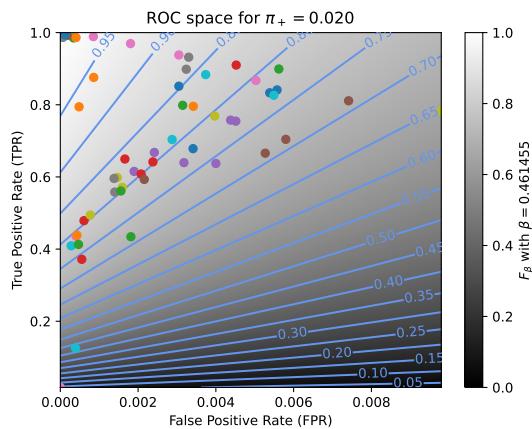


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

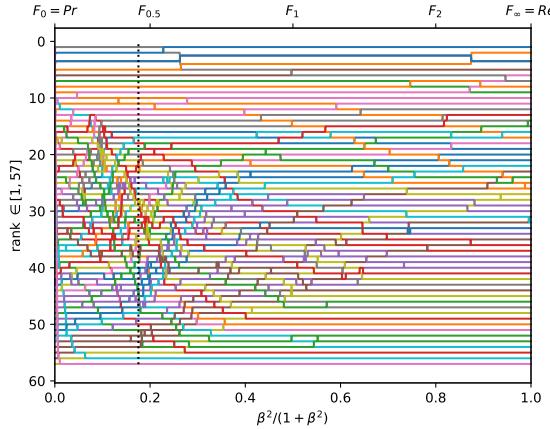


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

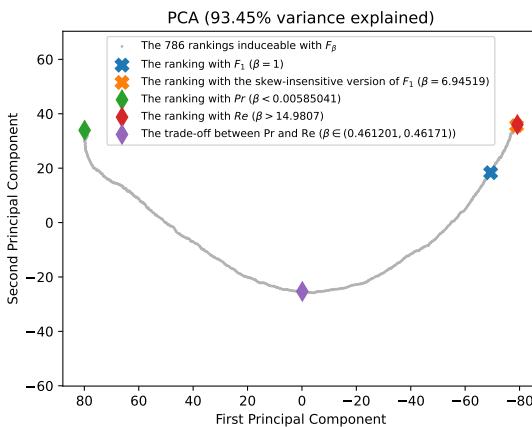
Figure A.3.36. Ranking of 57 BGS methods evaluated on the video "diningRoom" ($\pi_+ = 0.0859$).



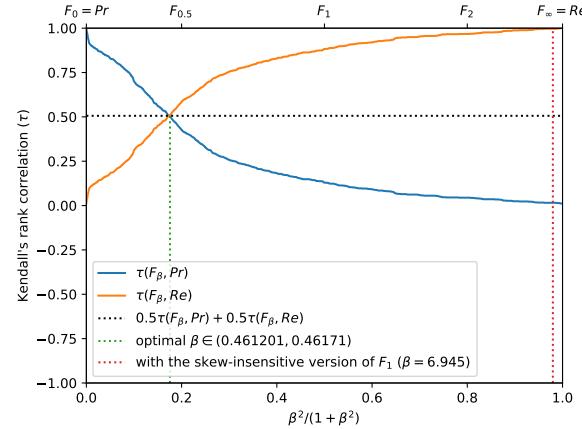
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



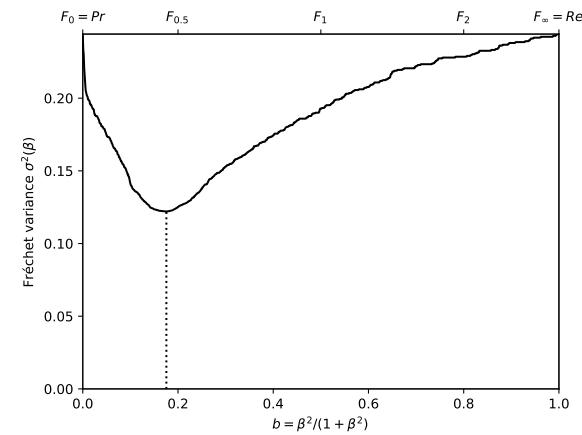
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



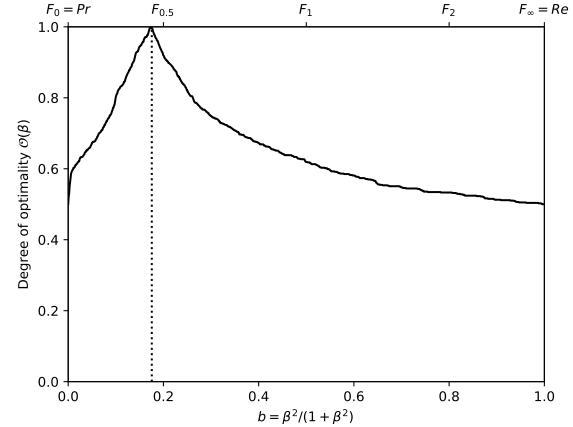
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

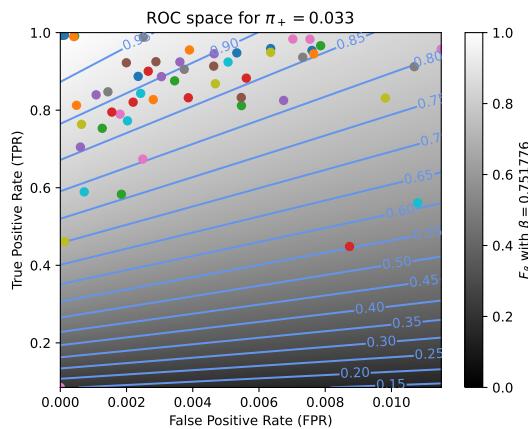


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

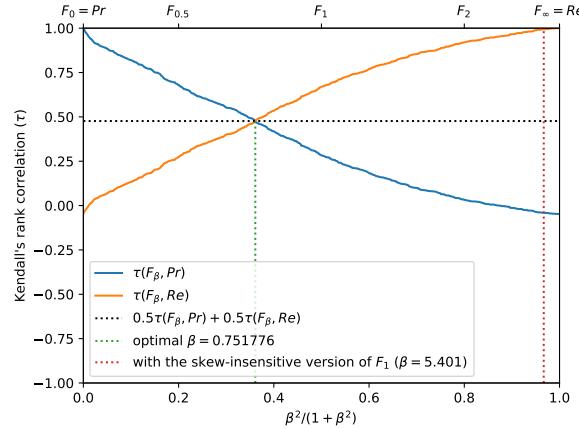


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

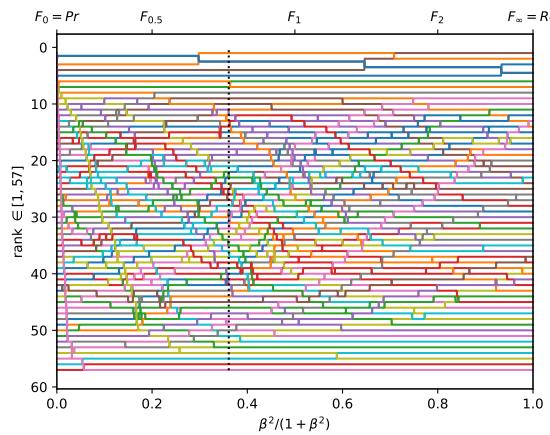
Figure A.3.37. Ranking of 57 BGS methods evaluated on the video "park" ($\pi_+ = 0.0203$).



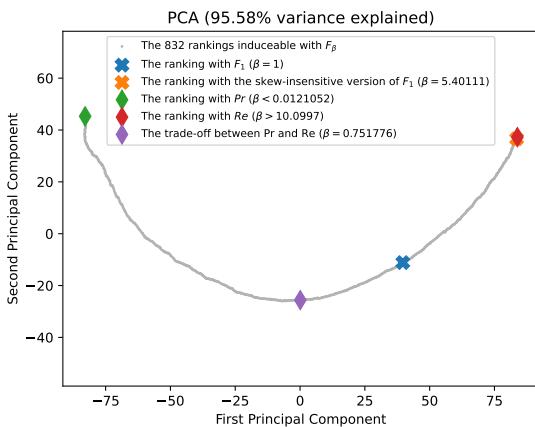
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



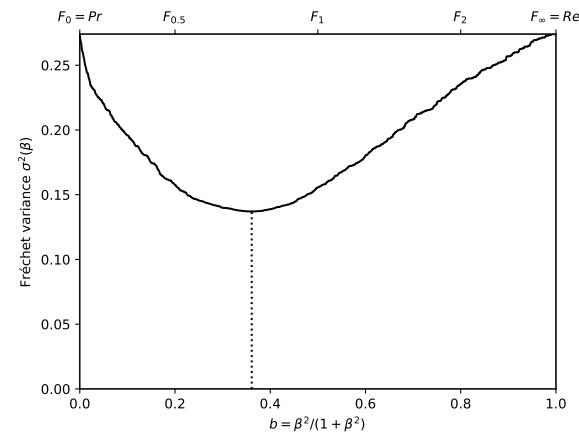
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



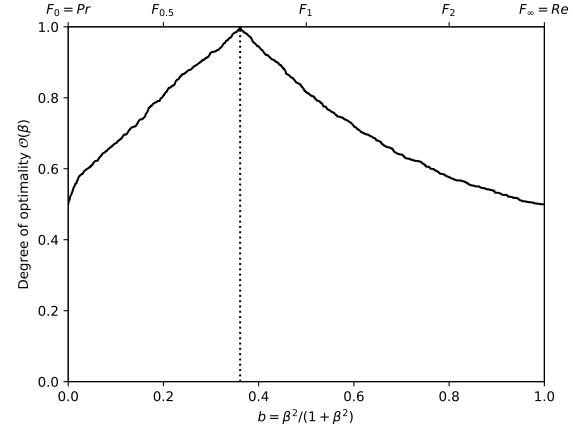
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

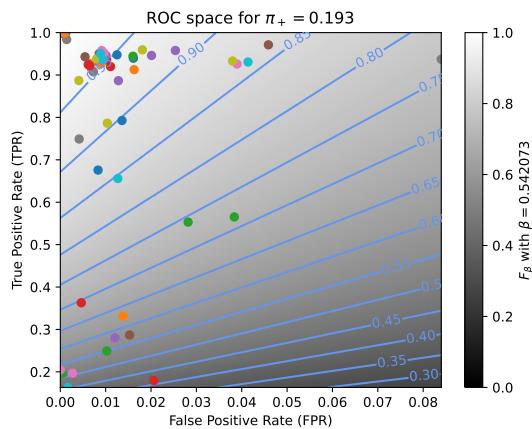


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

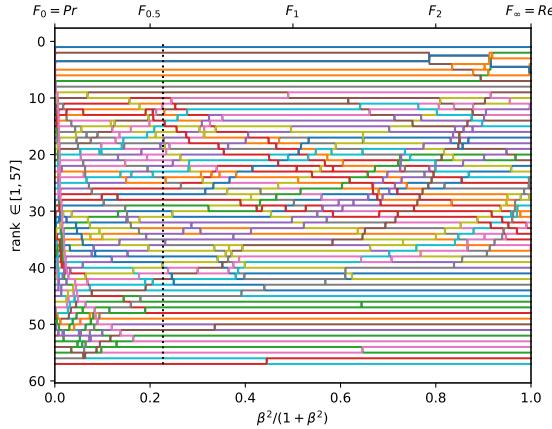


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

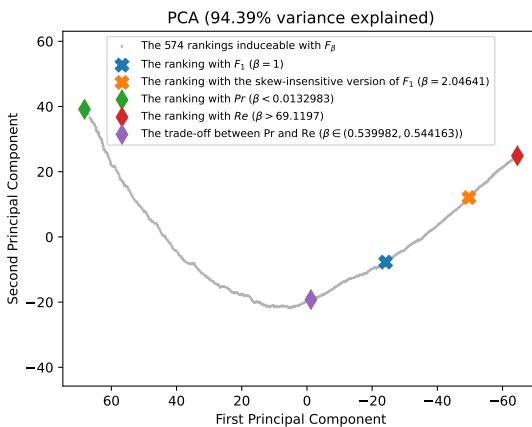
Figure A.3.38. Ranking of 57 BGS methods evaluated on the video "corridor" ($\pi_+ = 0.0331$).



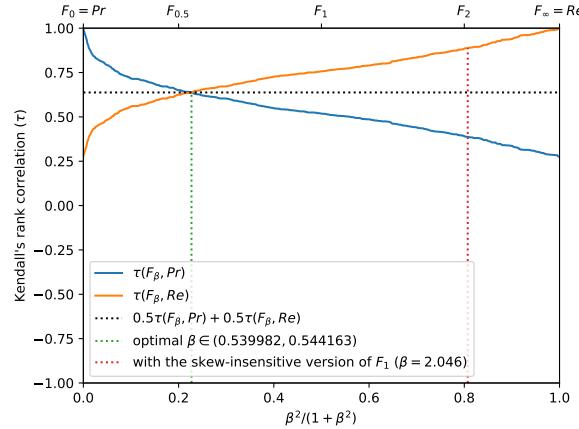
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



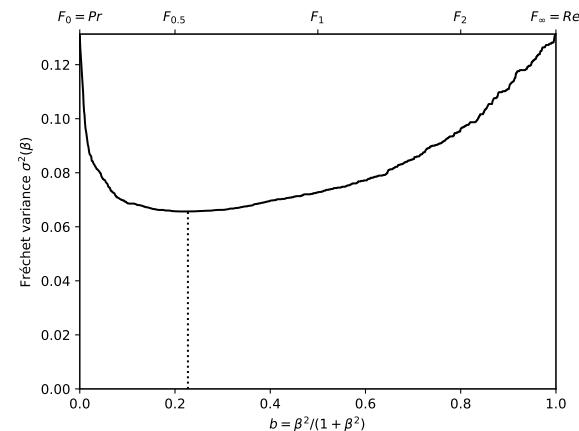
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



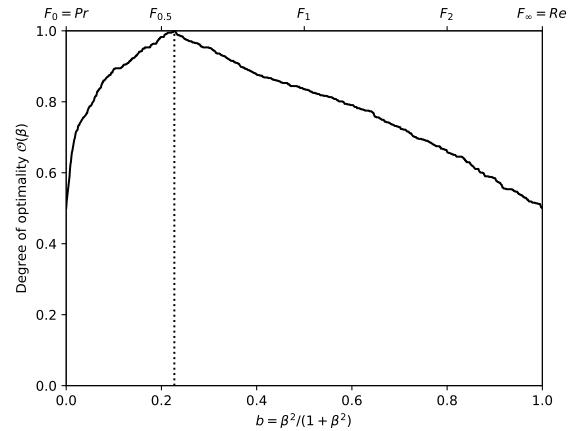
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

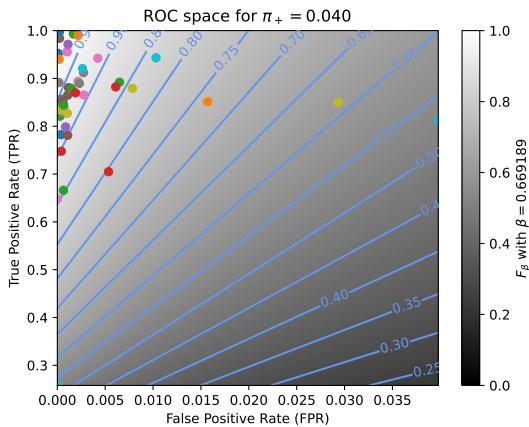


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

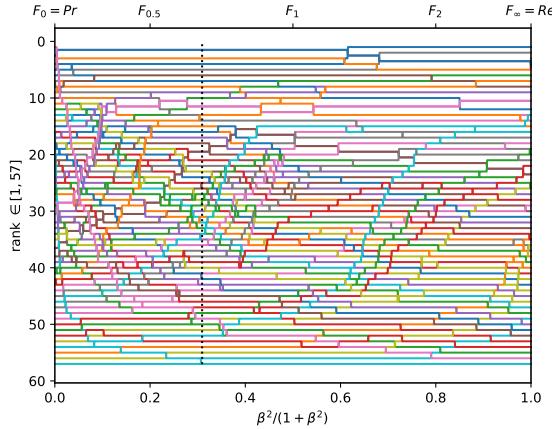


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

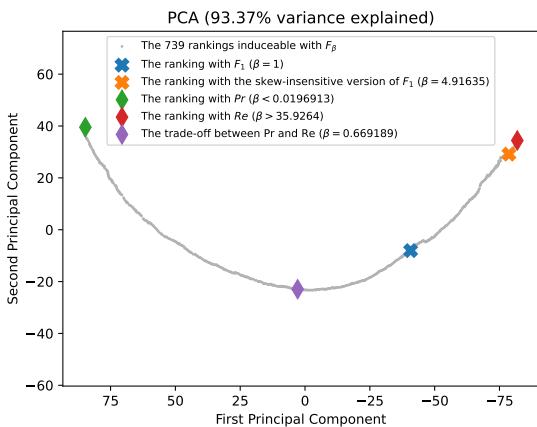
Figure A.3.39. Ranking of 57 BGS methods evaluated on the video "library" ($\pi_+ = 0.1928$).



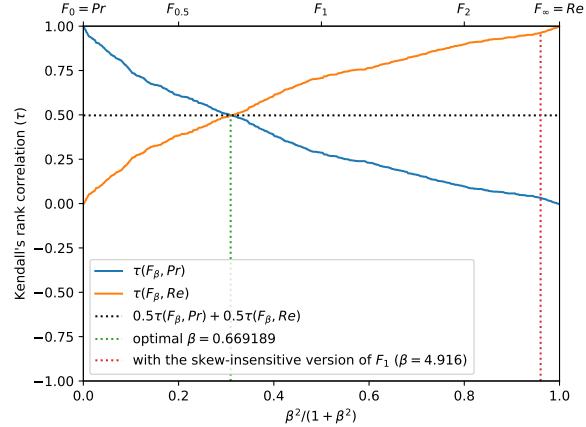
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



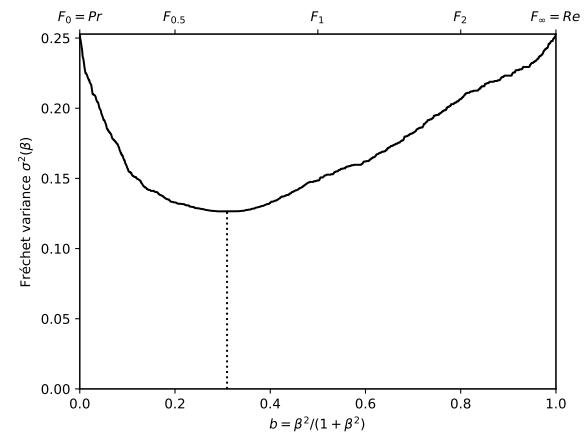
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



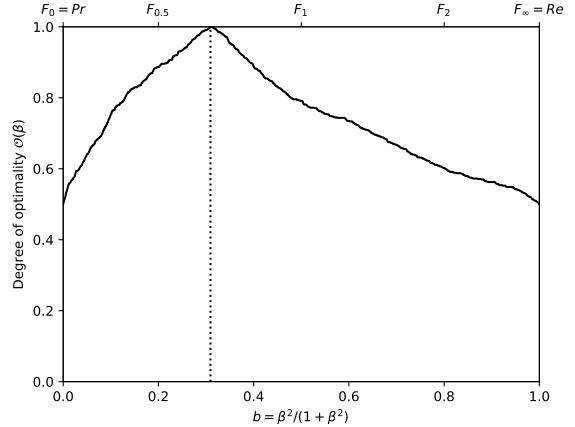
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

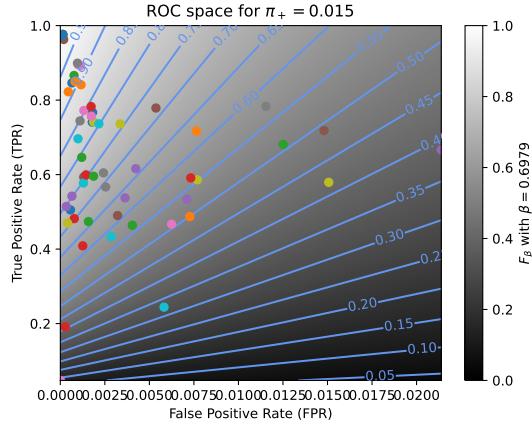


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

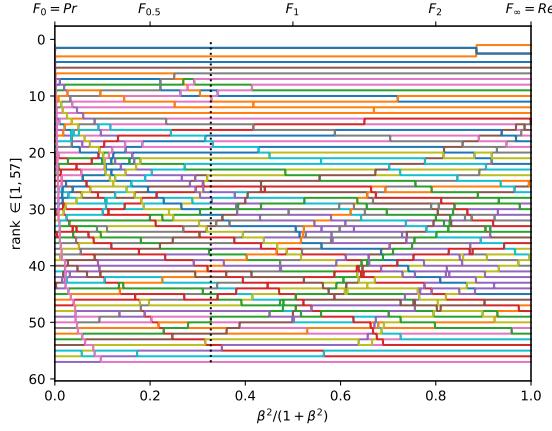


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

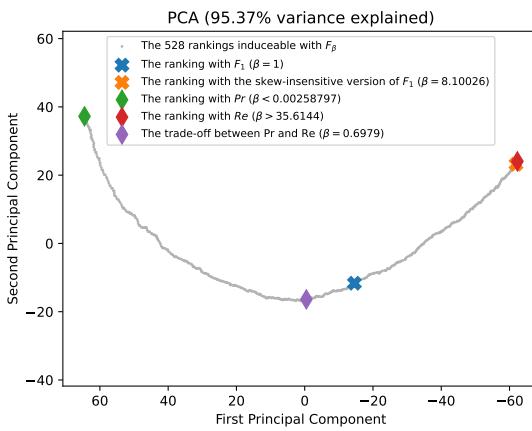
Figure A.3.40. Ranking of 57 BGS methods evaluated on the video "skating" ($\pi_+ = 0.0397$).



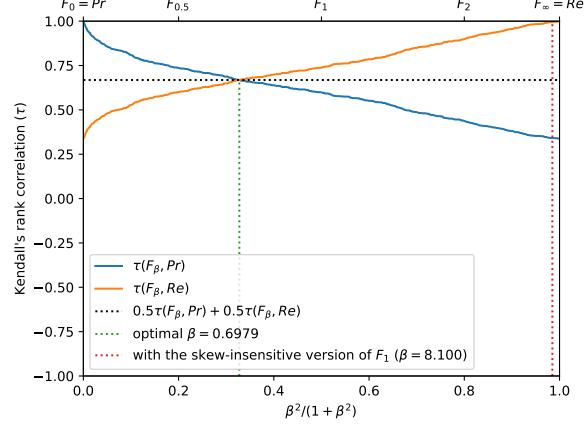
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



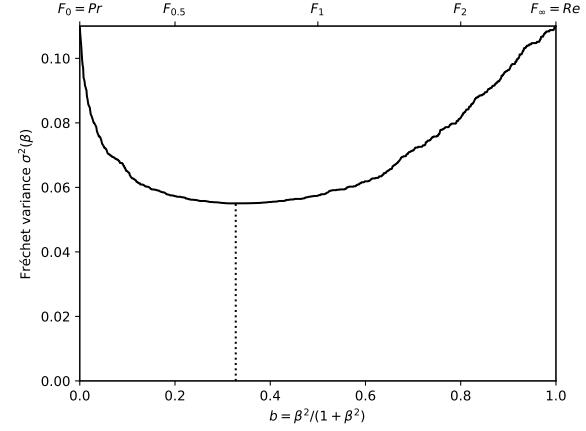
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



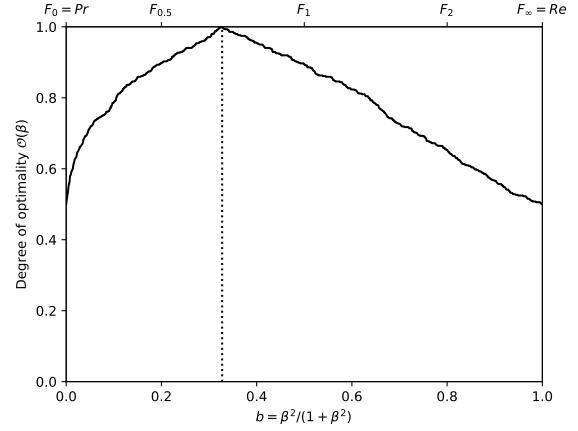
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



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

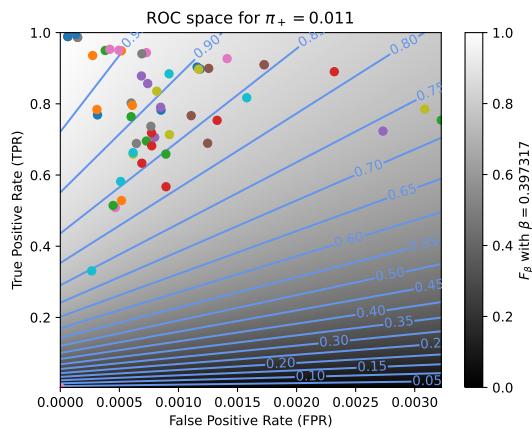


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

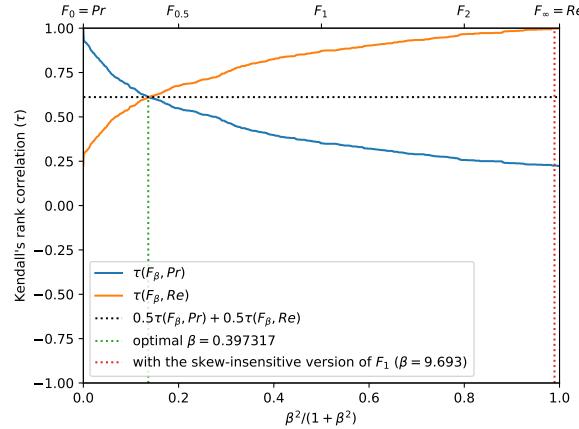


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

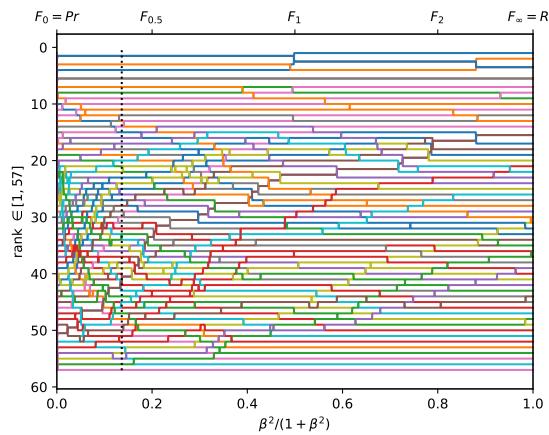
Figure A.3.41. Ranking of 57 BGS methods evaluated on the video "wetSnow" ($\pi_+ = 0.0150$).



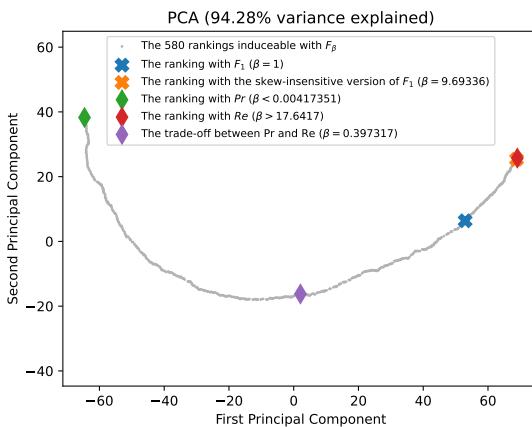
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



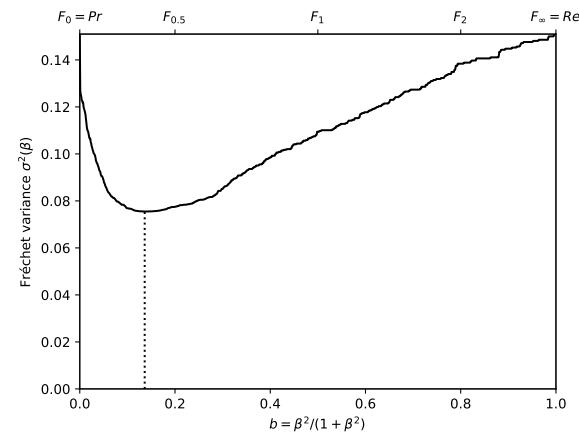
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



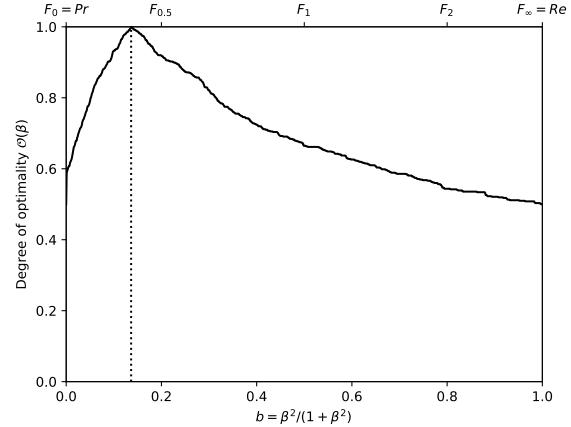
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

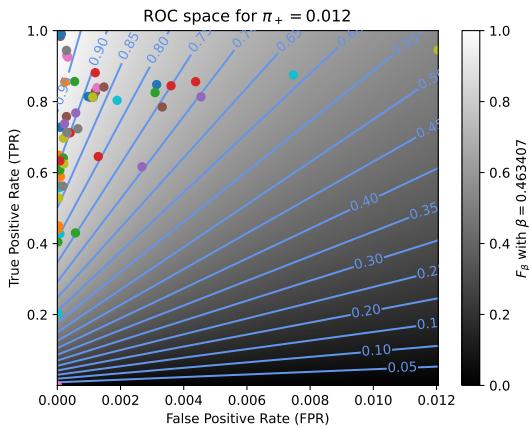


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

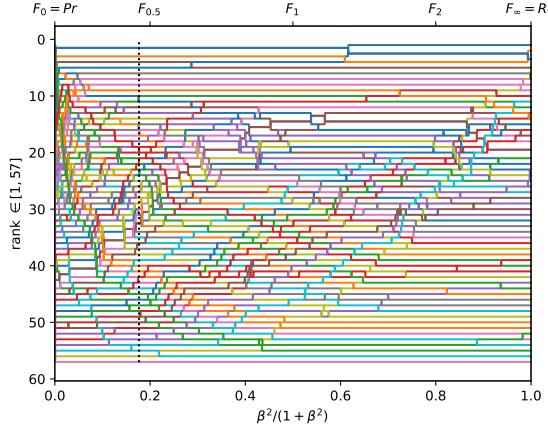


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

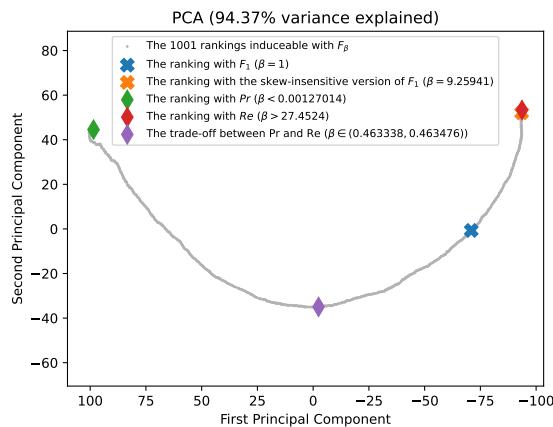
Figure A.3.42. Ranking of 57 BGS methods evaluated on the video "snowFall" ($\pi_+ = 0.0105$).



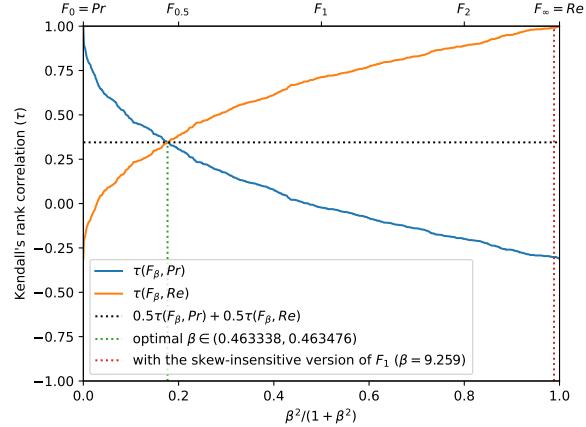
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



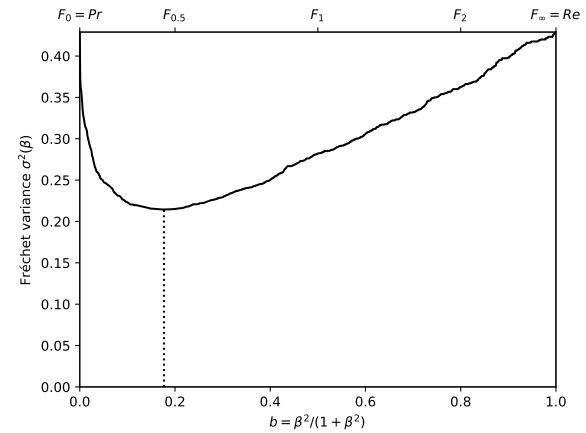
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



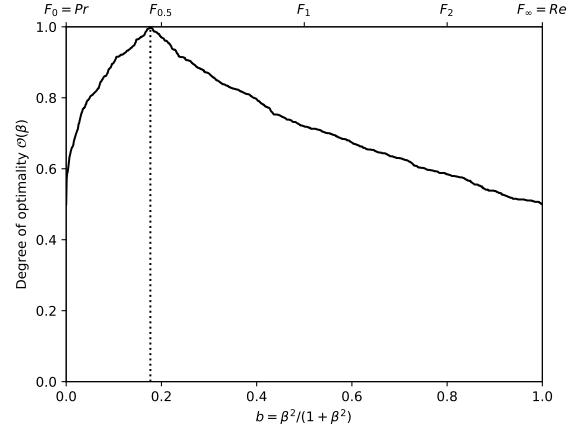
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

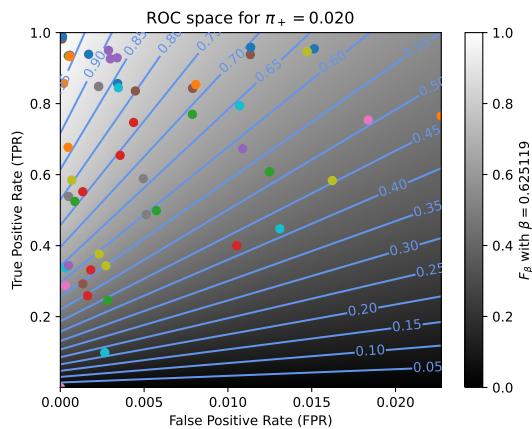


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

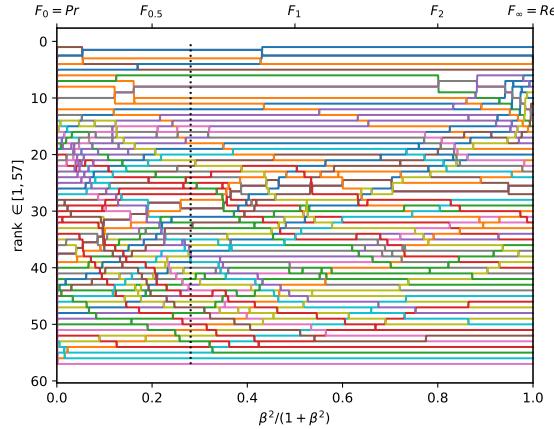


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

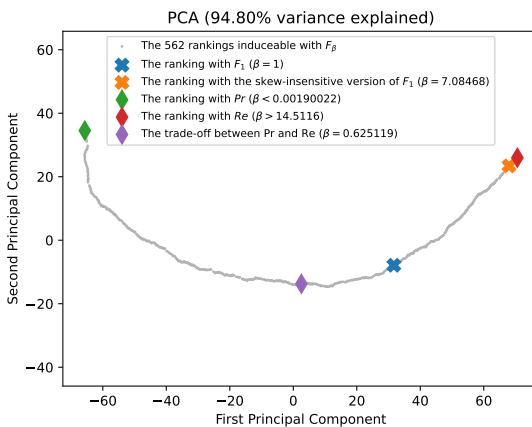
Figure A.3.43. Ranking of 57 BGS methods evaluated on the video "blizzard" ($\pi_+ = 0.0115$).



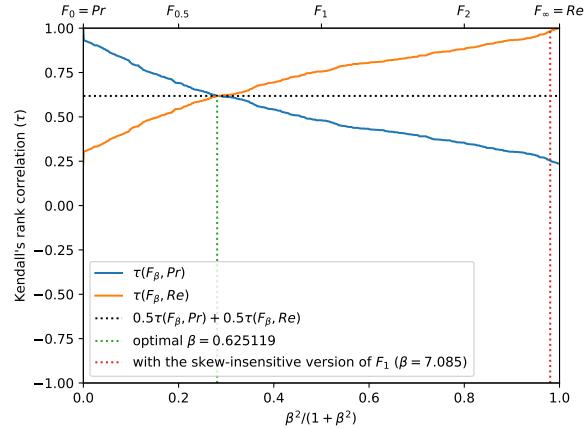
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



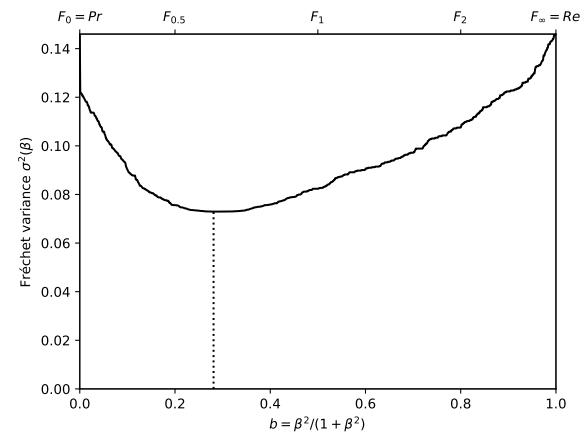
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



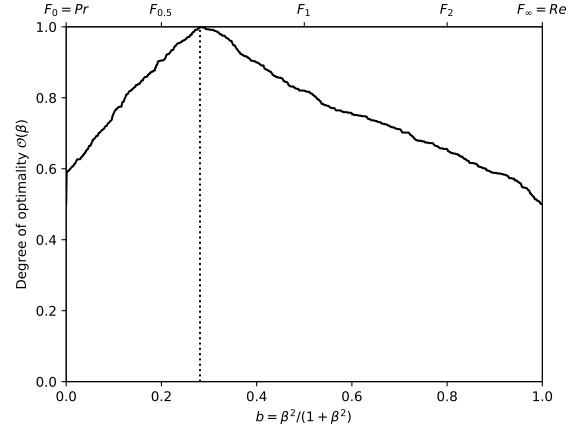
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

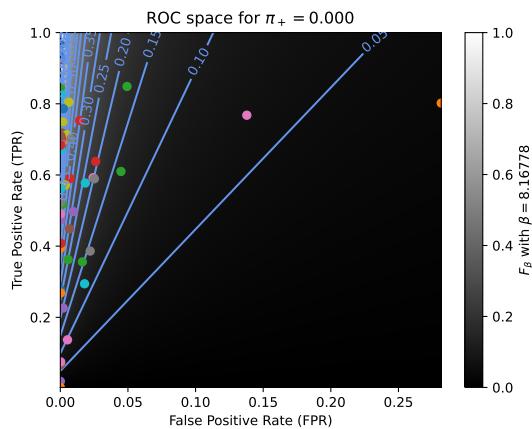


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

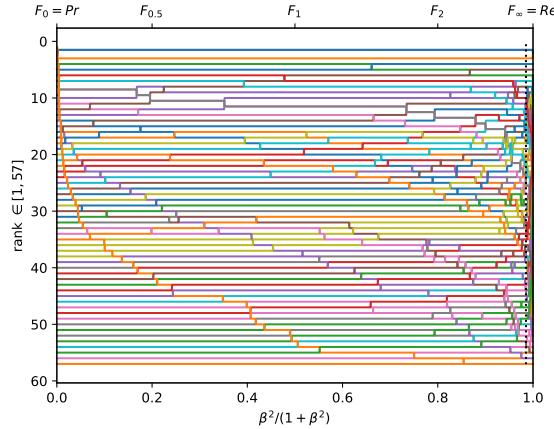


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

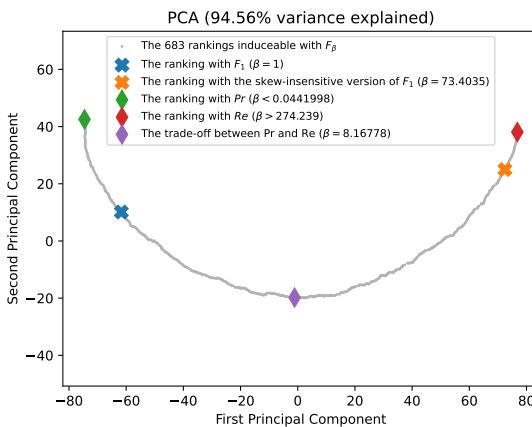
Figure A.3.44. Ranking of 57 BGS methods evaluated on the video "tunnelExit_0_35fps" ($\pi_+ = 0.0195$).



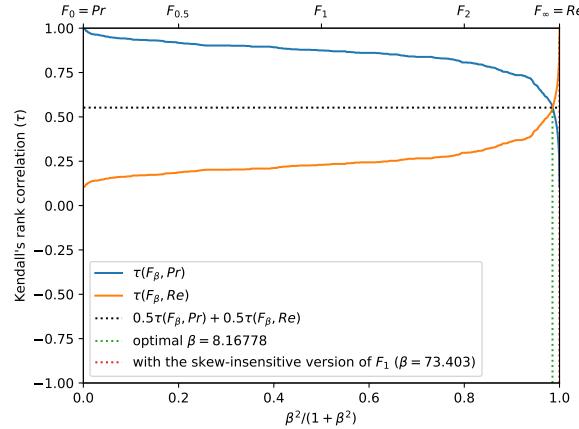
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



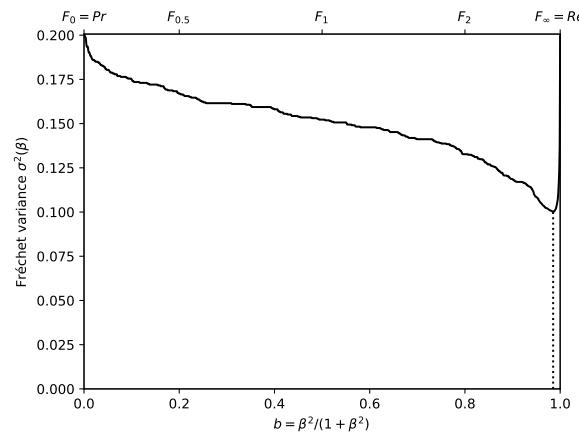
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



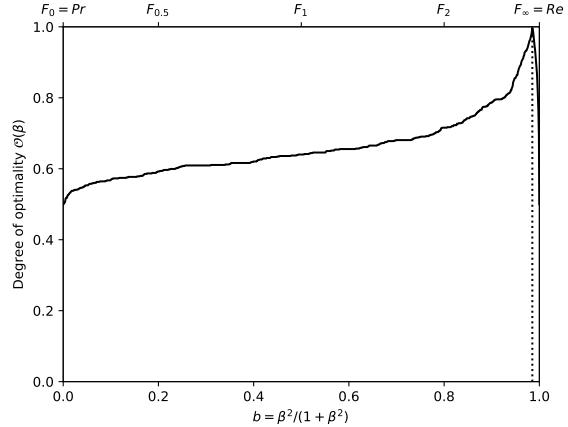
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

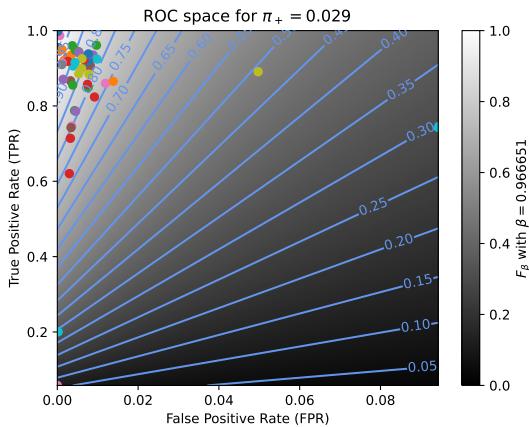


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

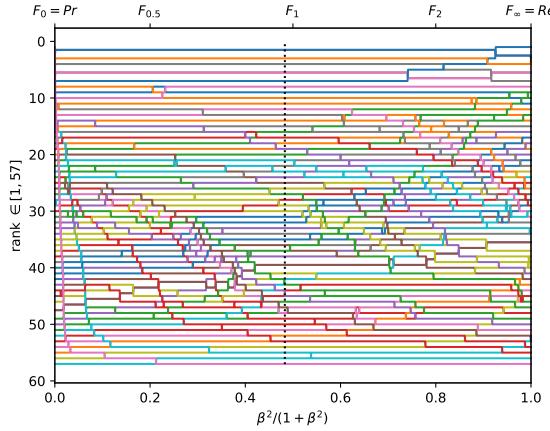


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

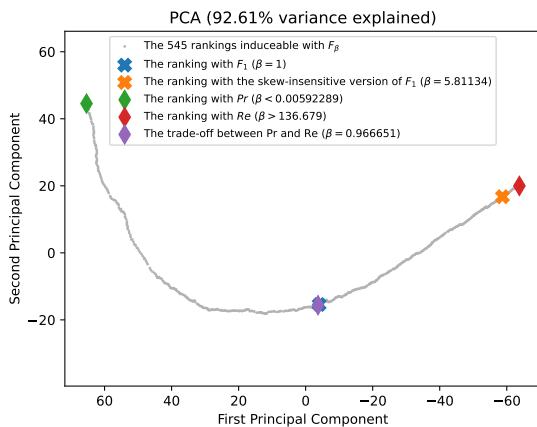
Figure A.3.45. Ranking of 57 BGS methods evaluated on the video "port_0_17fps" ($\pi_+ = 0.0002$).



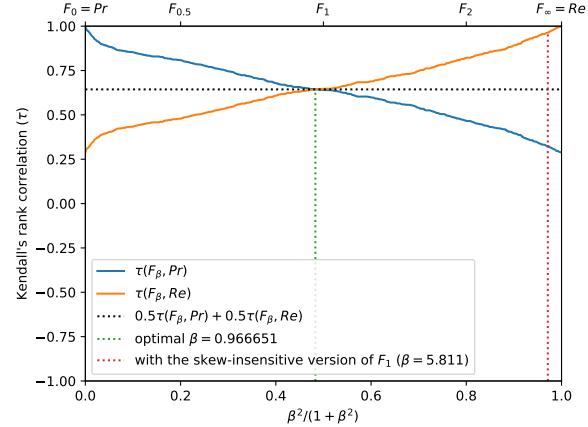
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



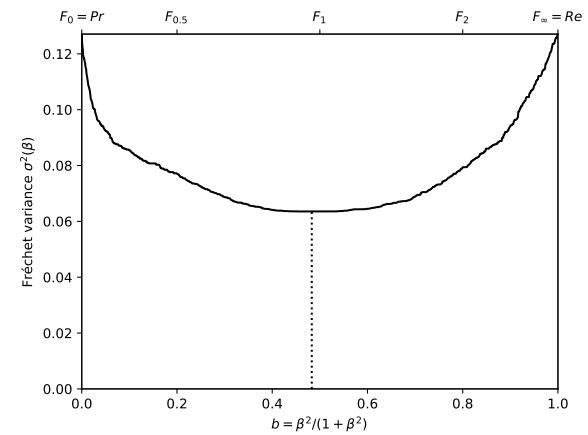
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



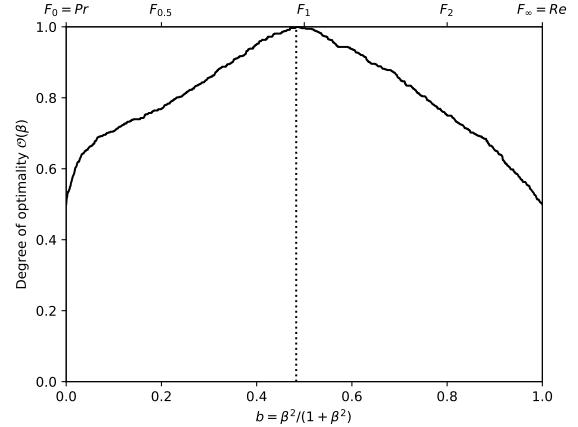
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

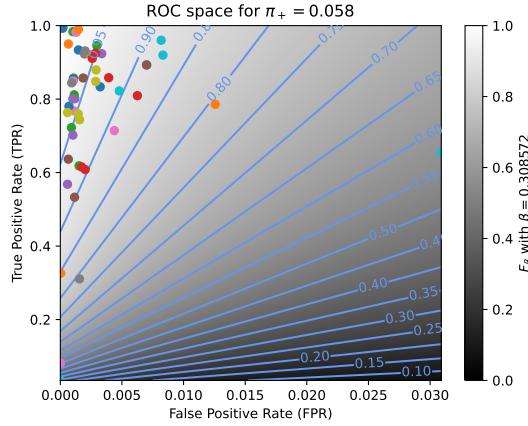


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

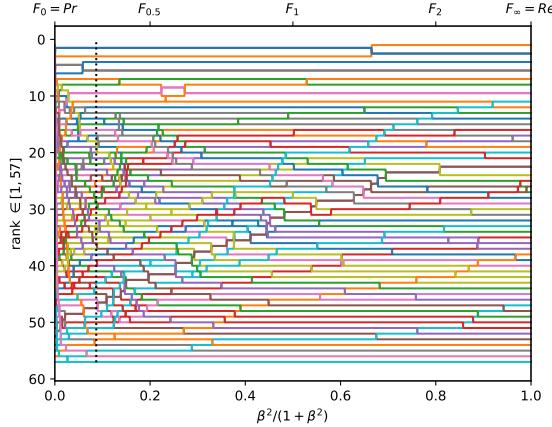


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

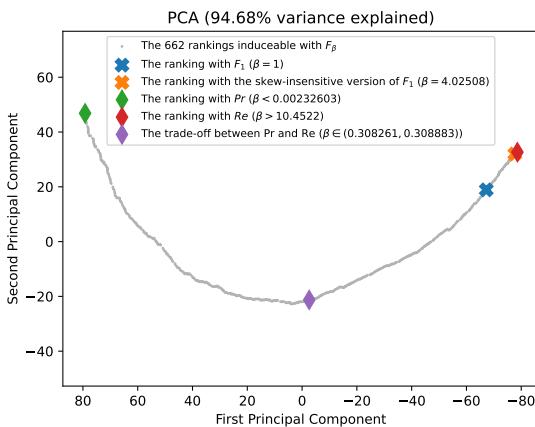
Figure A.3.46. Ranking of 57 BGS methods evaluated on the video "tramCrossroad_1fps" ($\pi_+ = 0.0288$).



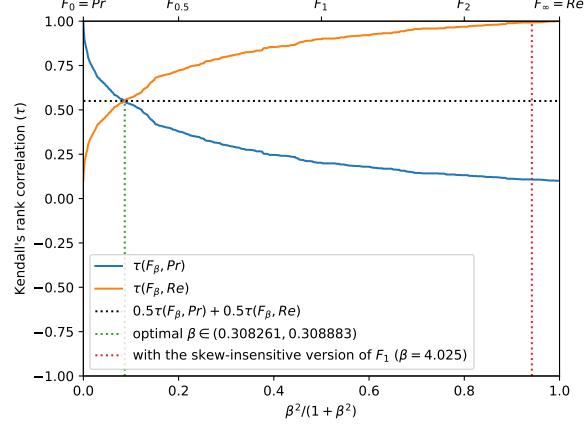
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



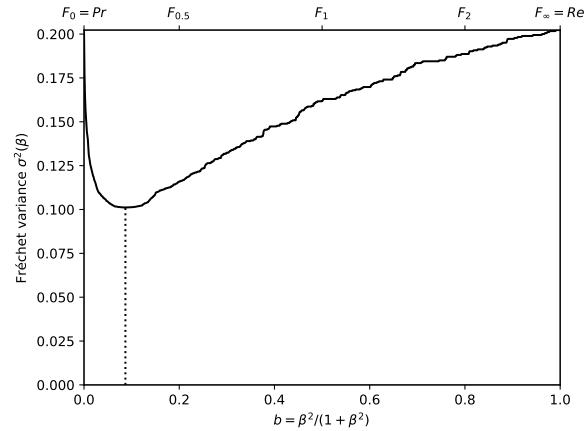
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



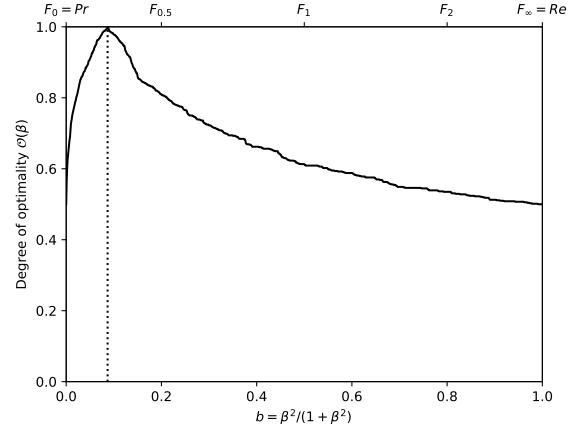
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

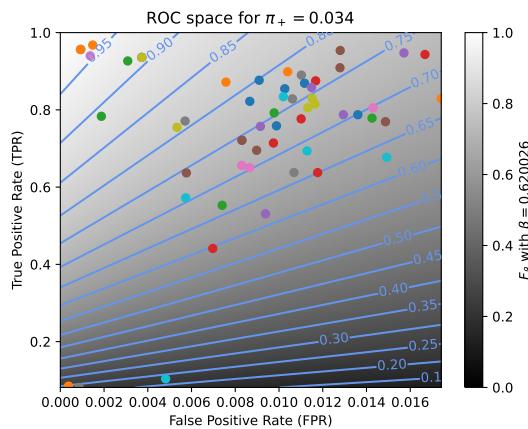


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

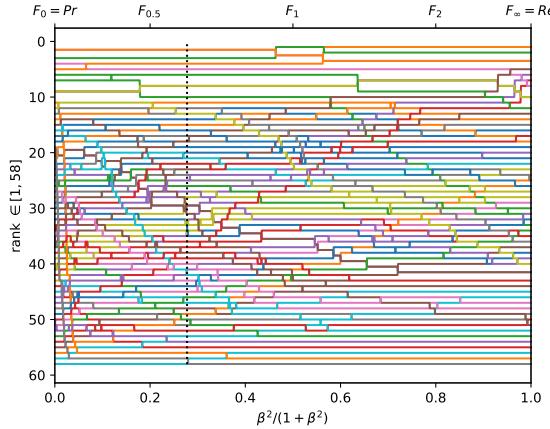


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

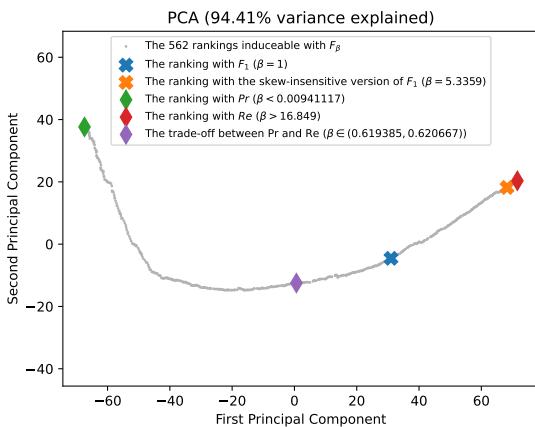
Figure A.3.47. Ranking of 57 BGS methods evaluated on the video "turnpike_0_5fps" ($\pi_+ = 0.0581$).



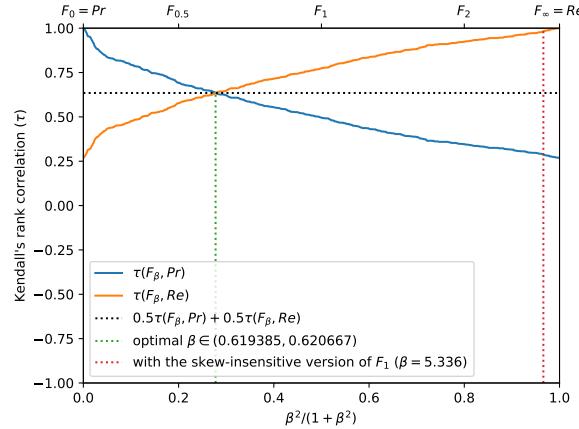
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



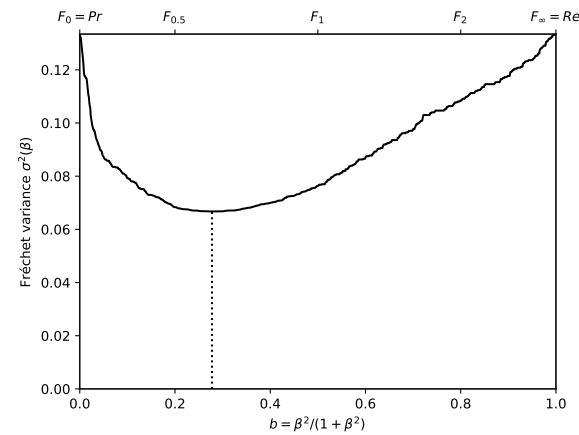
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



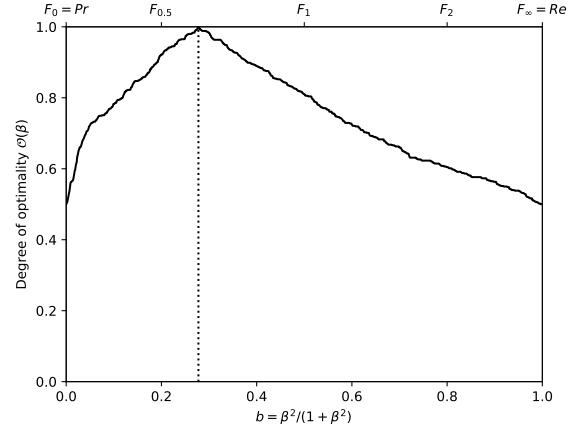
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

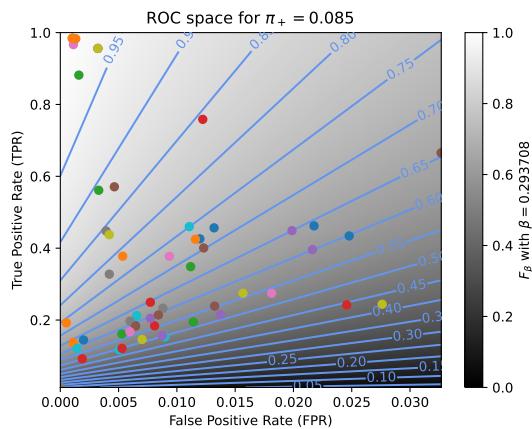


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

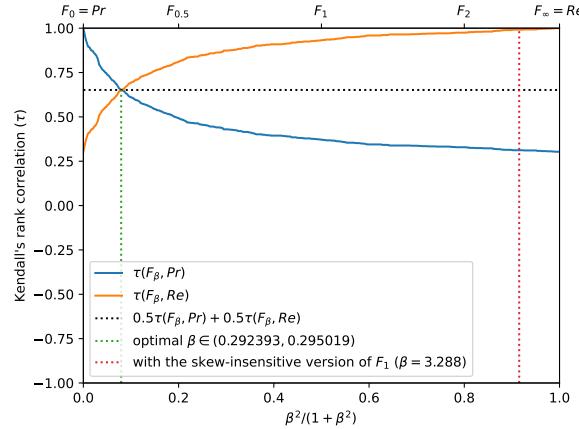


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

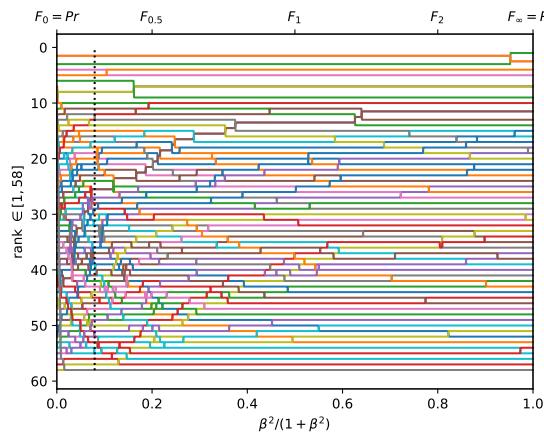
Figure A.3.48. Ranking of 58 BGS methods evaluated on the video "tramStation" ($\pi_+ = 0.0339$).



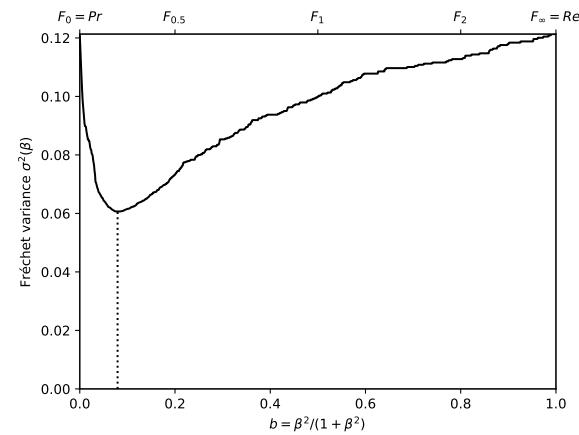
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



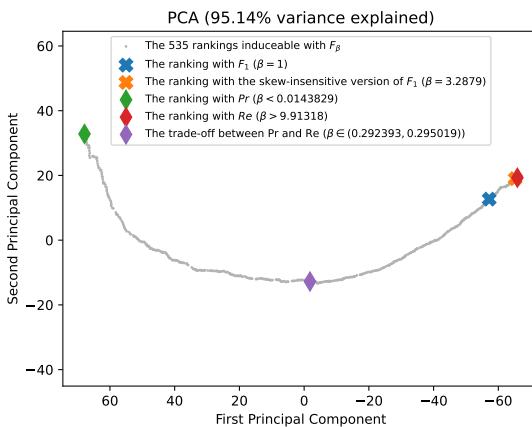
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



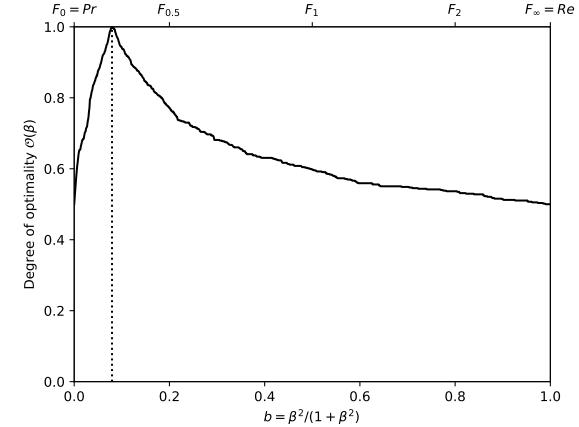
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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.

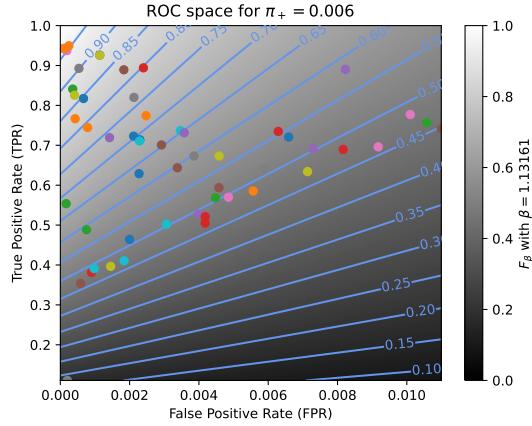


(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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

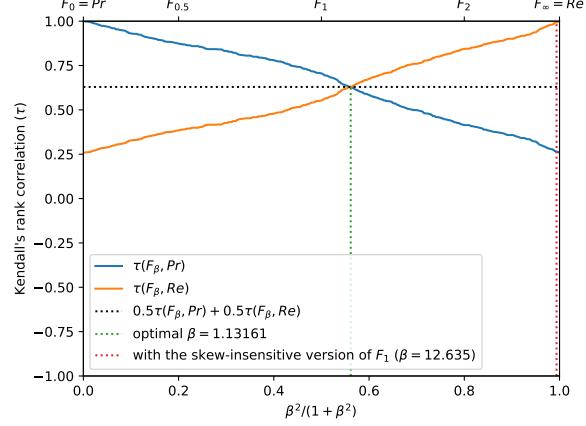


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

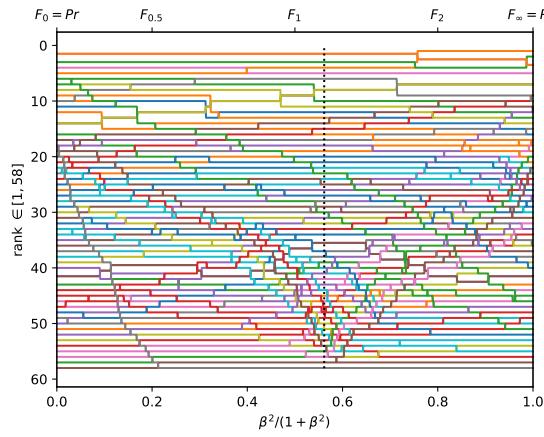
Figure A.3.49. Ranking of 58 BGS methods evaluated on the video "busyBoulevard" ($\pi_+ = 0.0847$).



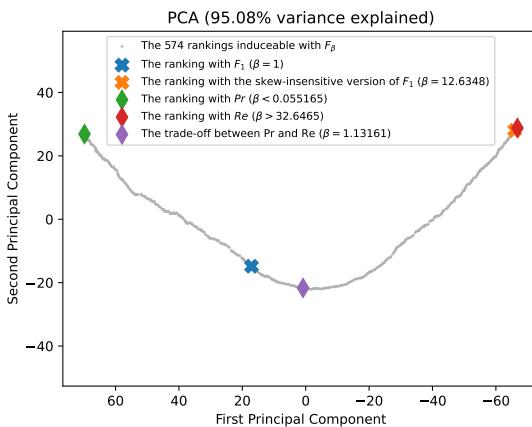
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



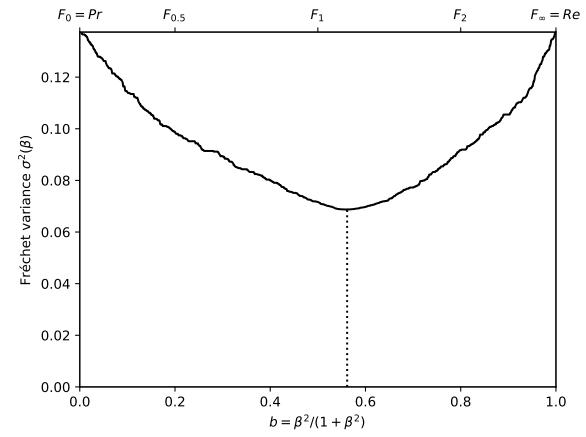
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



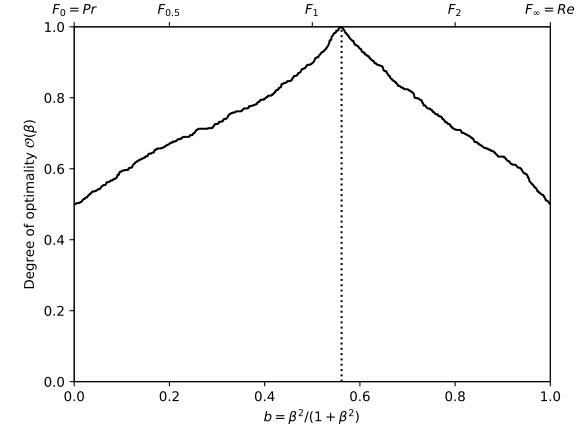
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

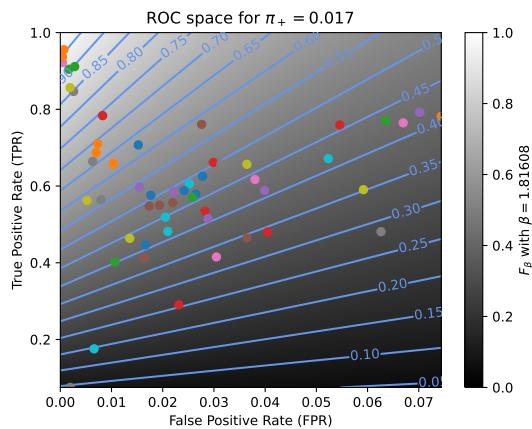


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

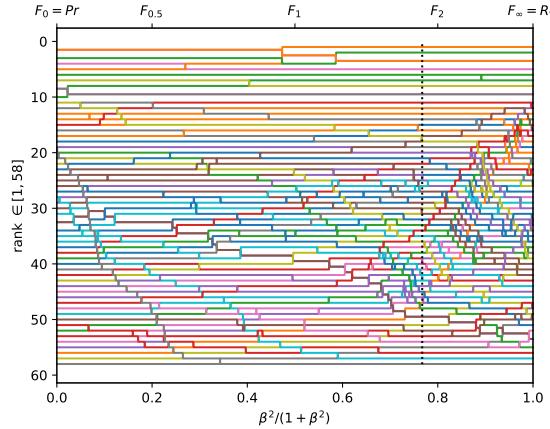


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

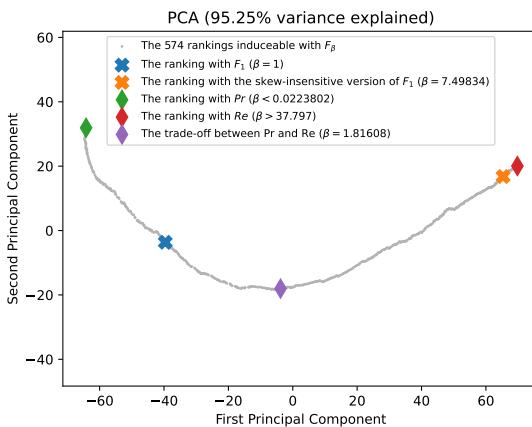
Figure A.3.50. Ranking of 58 BGS methods evaluated on the video "streetCornerAtNight" ($\pi_+ = 0.0062$).



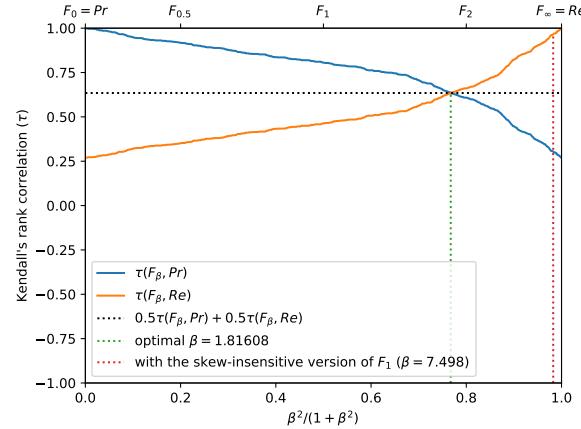
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



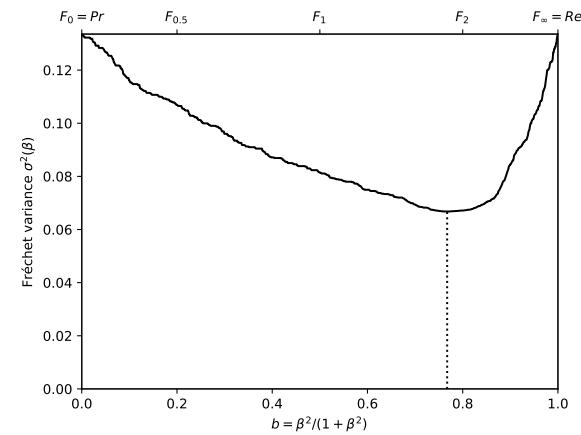
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



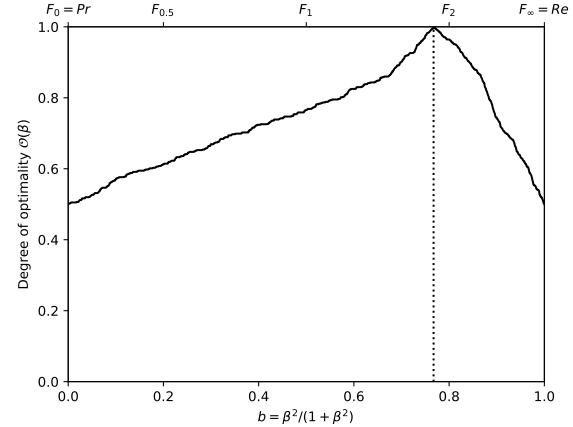
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



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

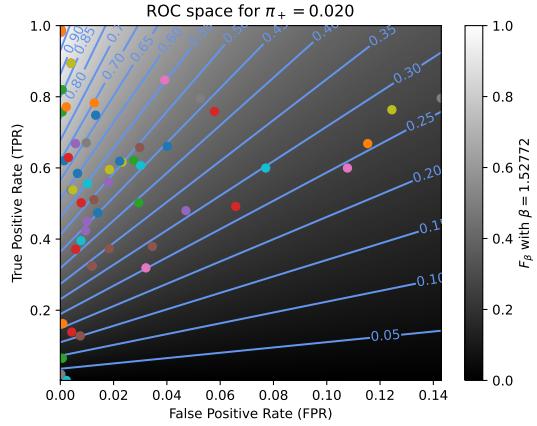


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(Re; F_\beta)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

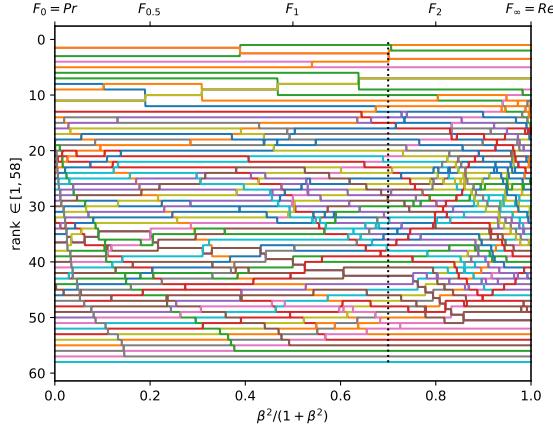


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

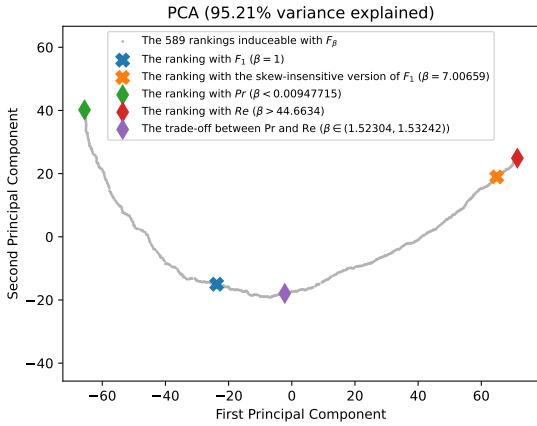
Figure A.3.51. Ranking of 58 BGS methods evaluated on the video "fluidHighway" ($\pi_+ = 0.0175$).



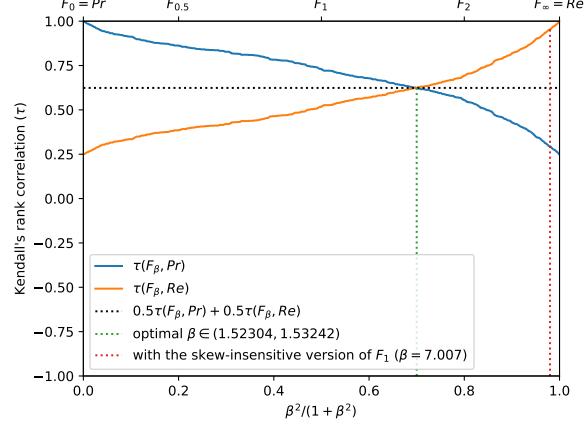
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



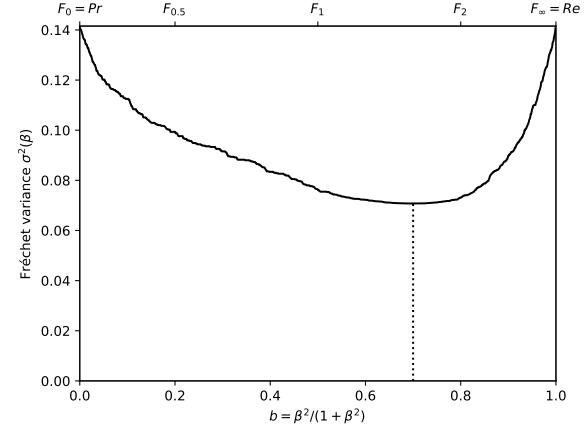
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



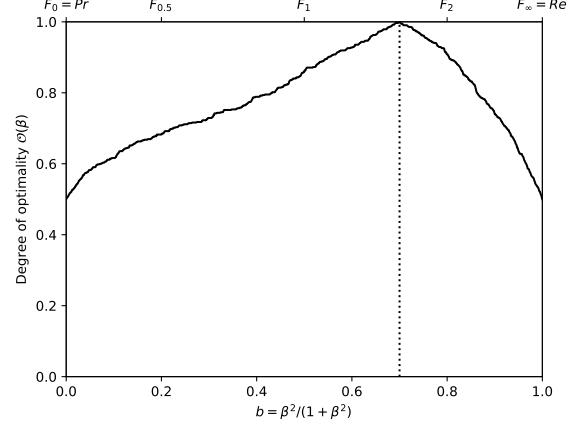
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

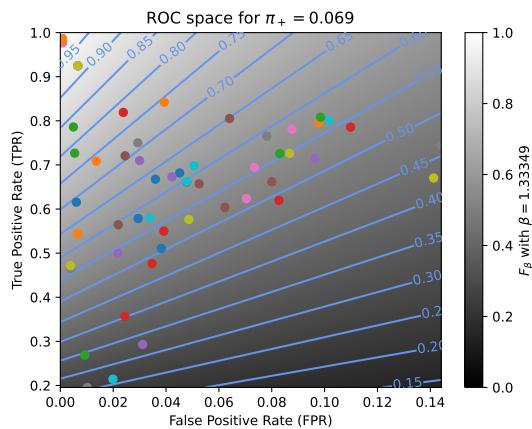


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

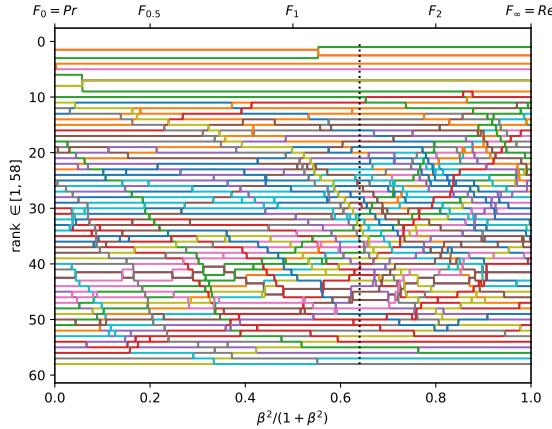


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

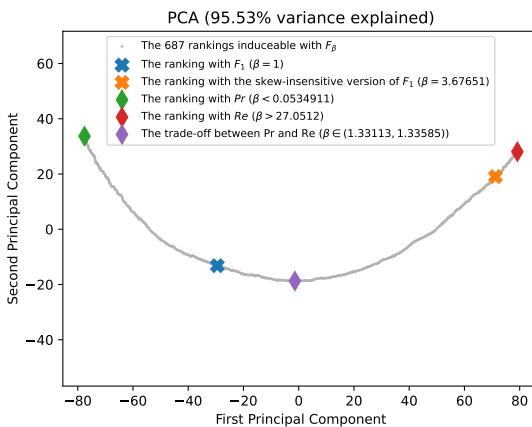
Figure A.3.52. Ranking of 58 BGS methods evaluated on the video "bridgeEntry" ($\pi_+ = 0.0200$).



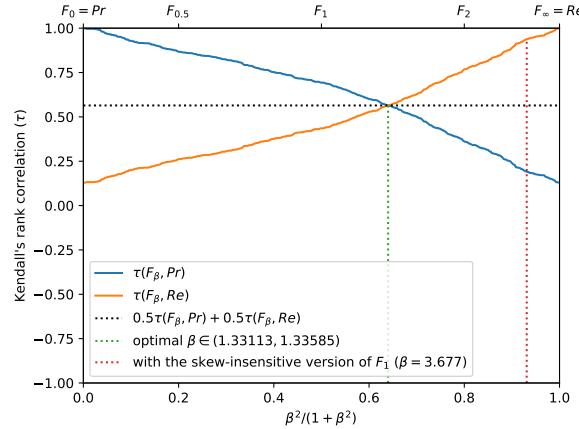
(a) The performances of 58 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



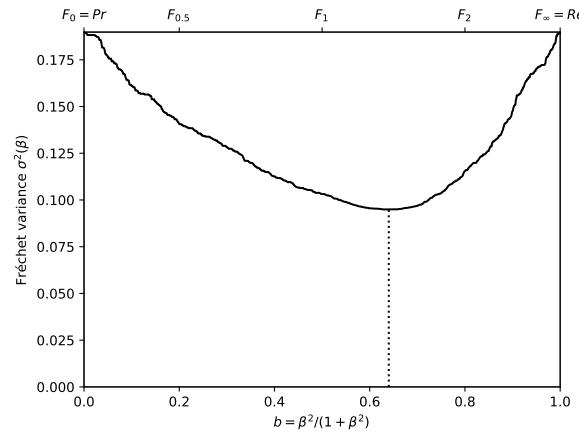
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



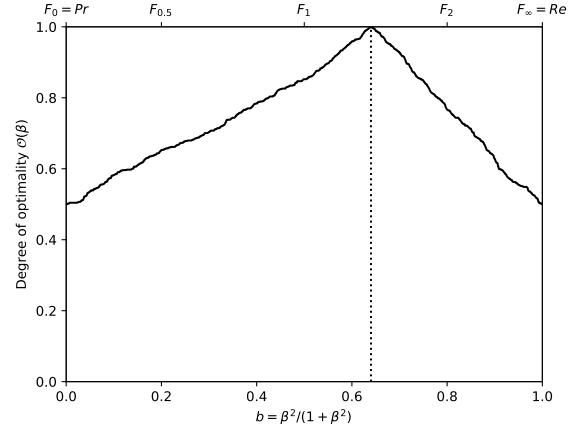
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

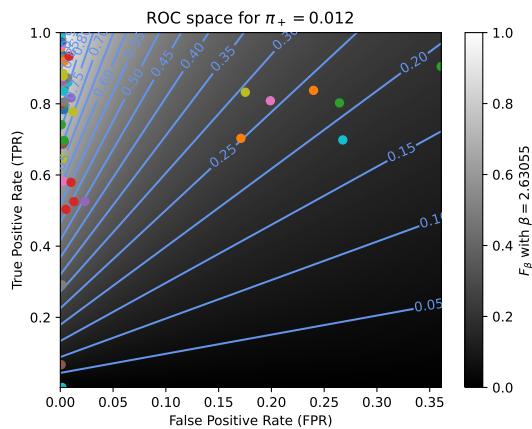


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

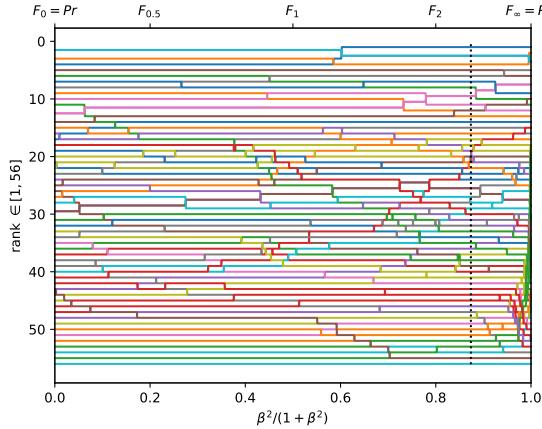


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

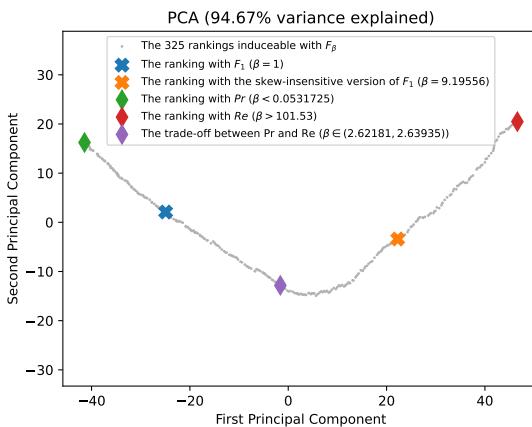
Figure A.3.53. Ranking of 58 BGS methods evaluated on the video "winterStreet" ($\pi_+ = 0.0689$).



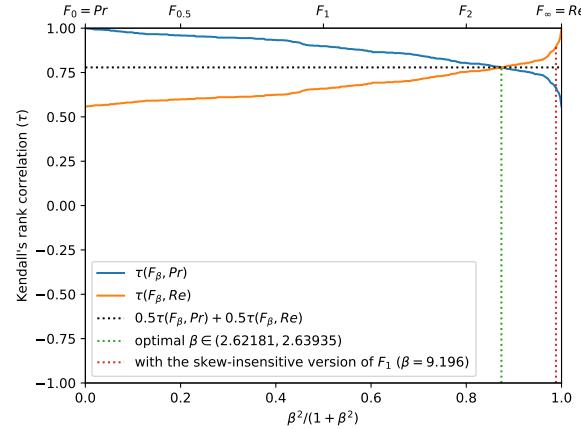
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



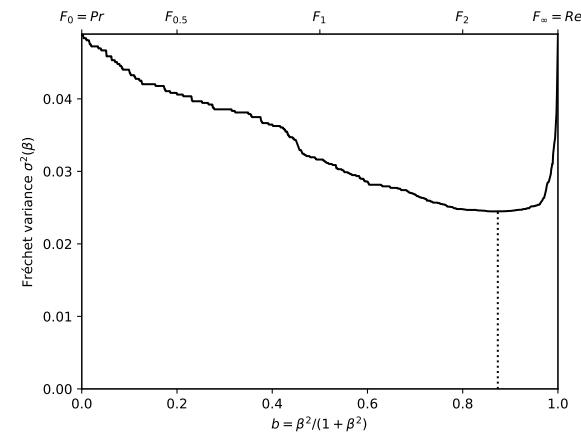
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



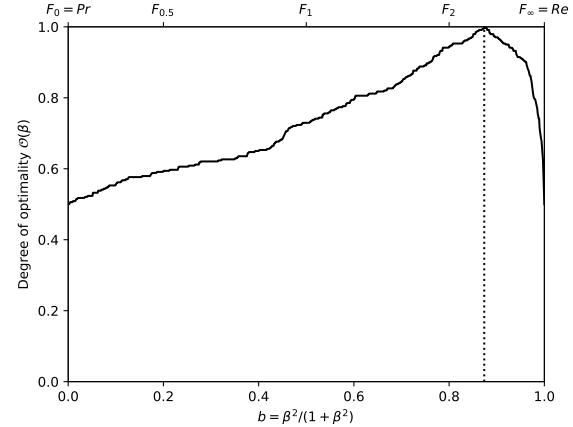
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

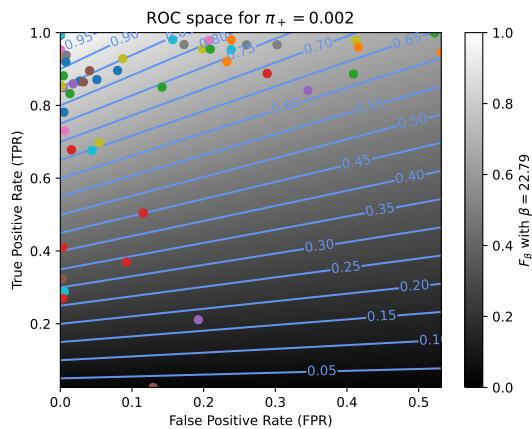


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

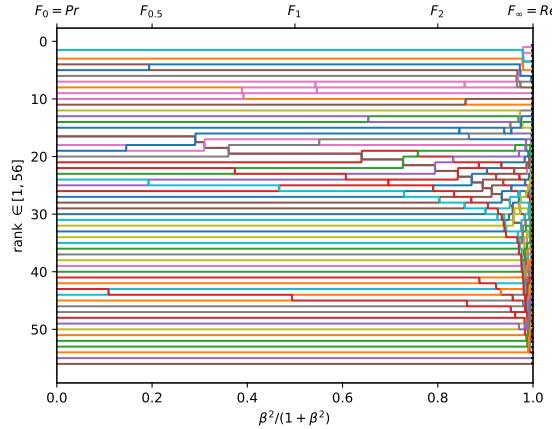


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

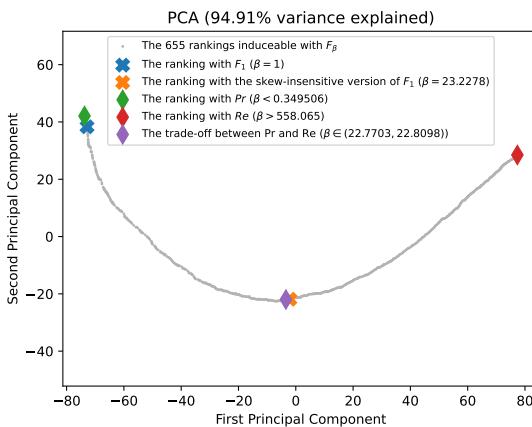
Figure A.3.54. Ranking of 56 BGS methods evaluated on the video "twoPositionPTZCam" ($\pi_+ = 0.0117$).



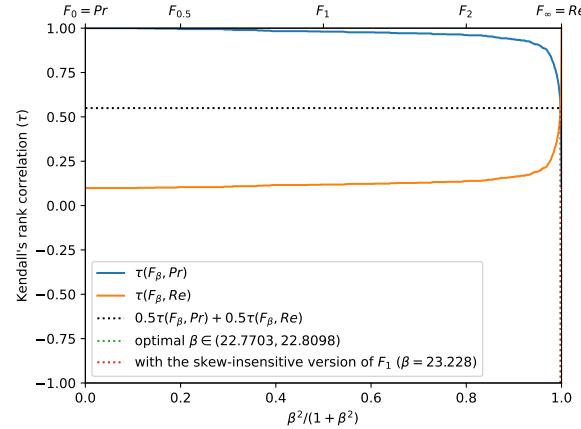
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



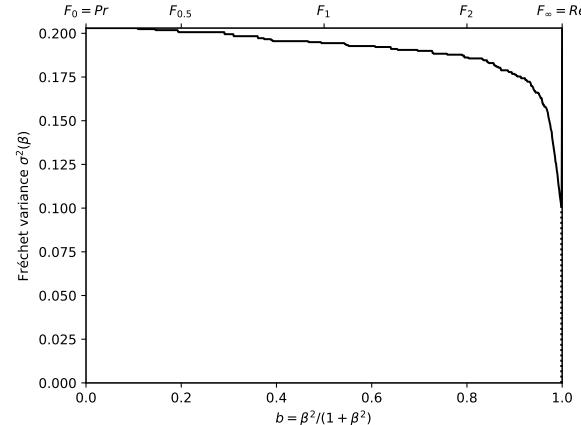
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



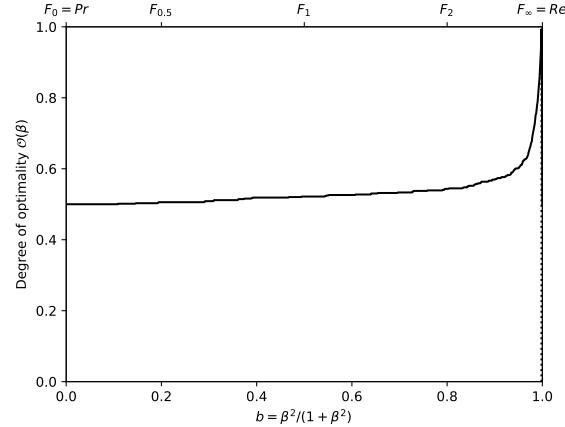
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

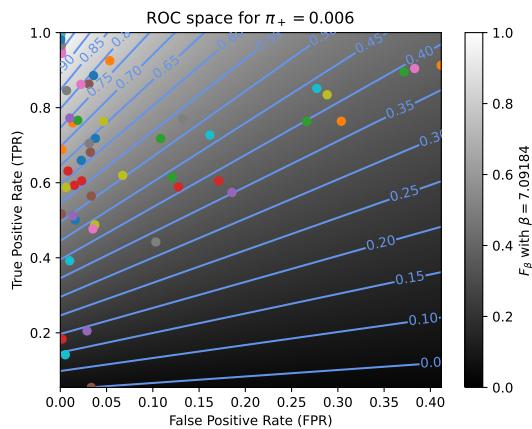


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

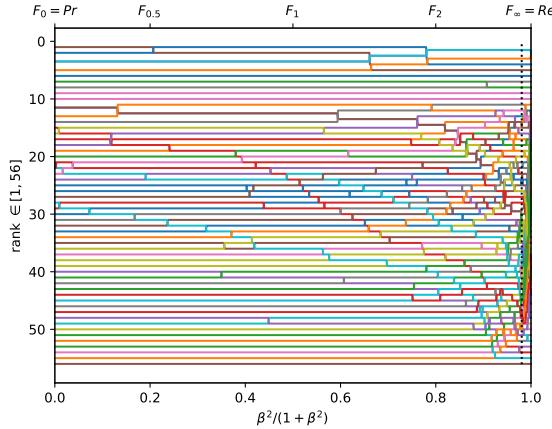


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

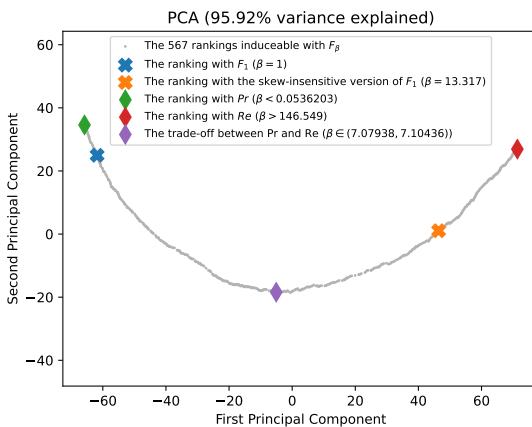
Figure A.3.55. Ranking of 56 BGS methods evaluated on the video "zoomInZoomOut" ($\pi_+ = 0.0019$).



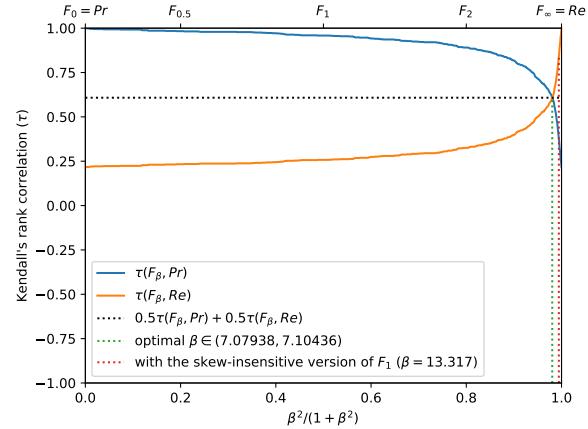
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



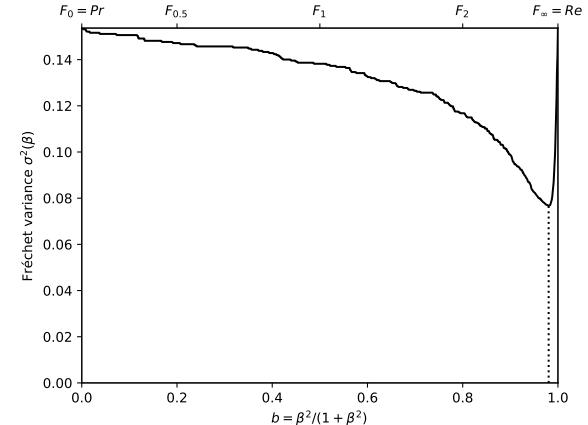
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



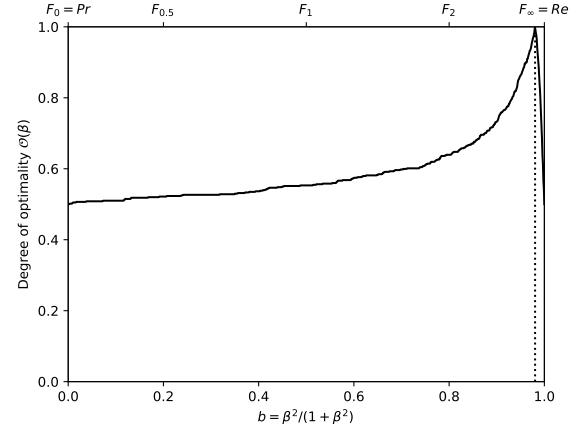
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

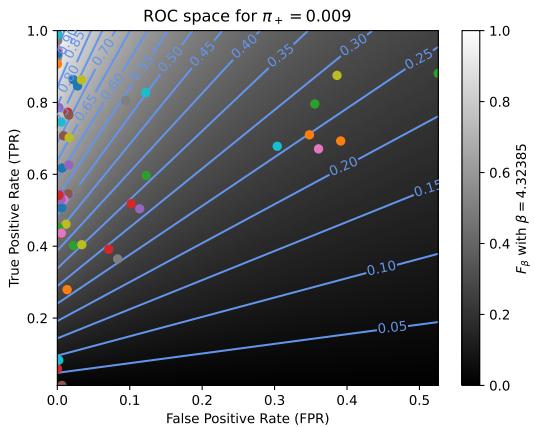


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

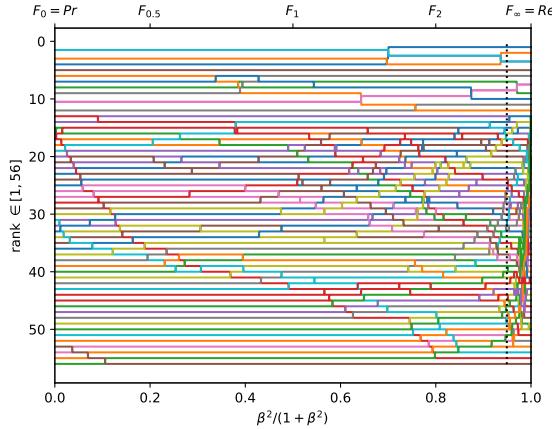


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

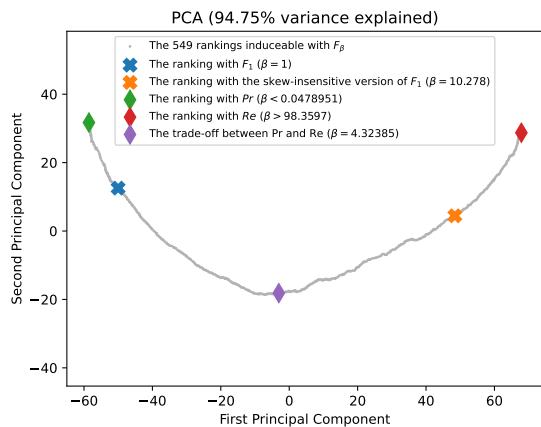
Figure A.3.56. Ranking of 56 BGS methods evaluated on the video "continuousPan" ($\pi_+ = 0.0056$).



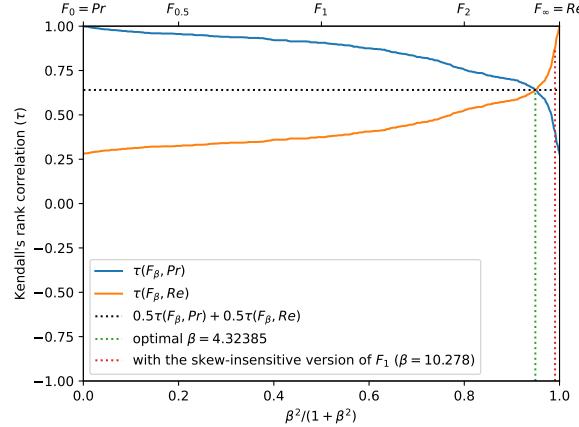
(a) The performances of 56 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



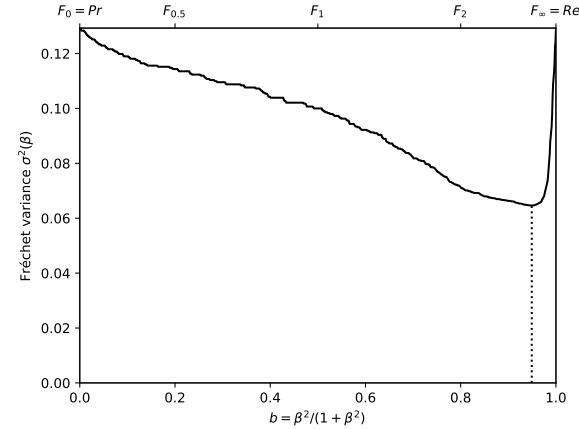
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



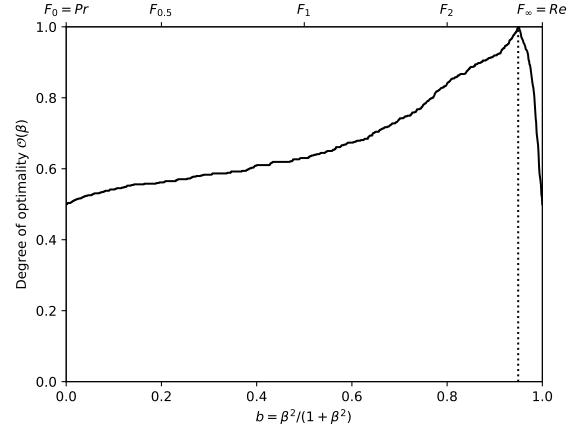
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

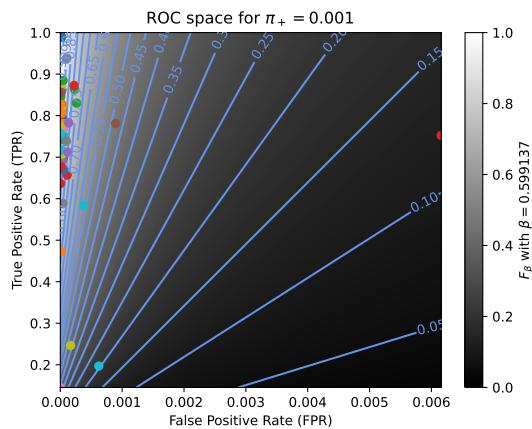


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

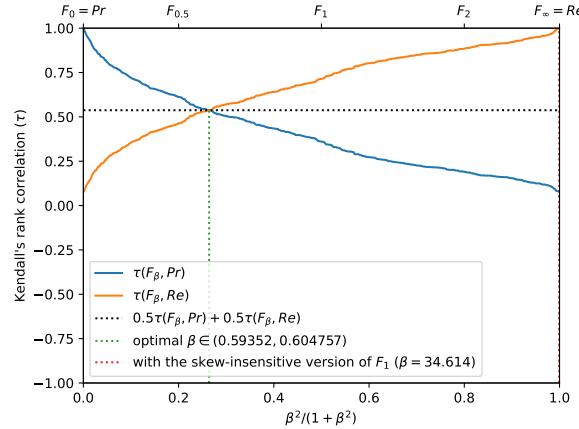


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

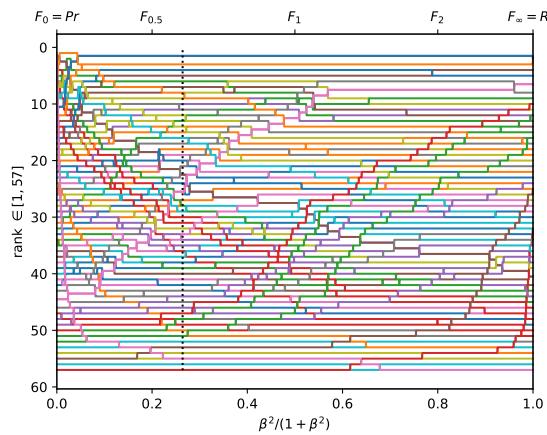
Figure A.3.57. Ranking of 56 BGS methods evaluated on the video "intermittentPan" ($\pi_+ = 0.0094$).



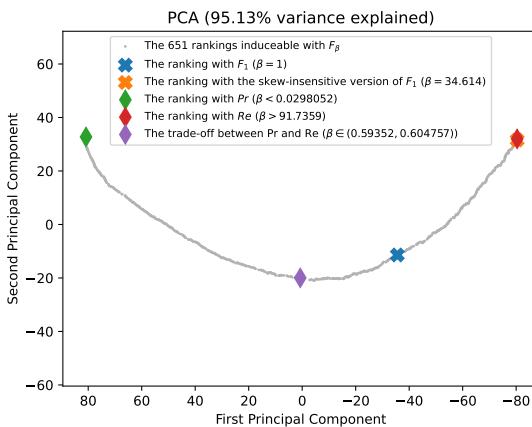
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



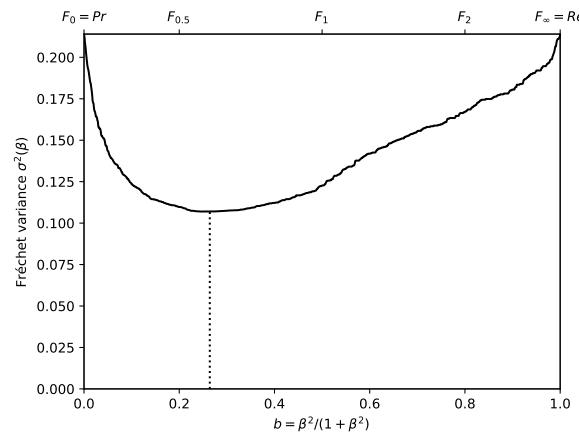
(b) The rank correlations $\tau(Pr; F_\beta)$ and $\tau(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the two curves intersect.



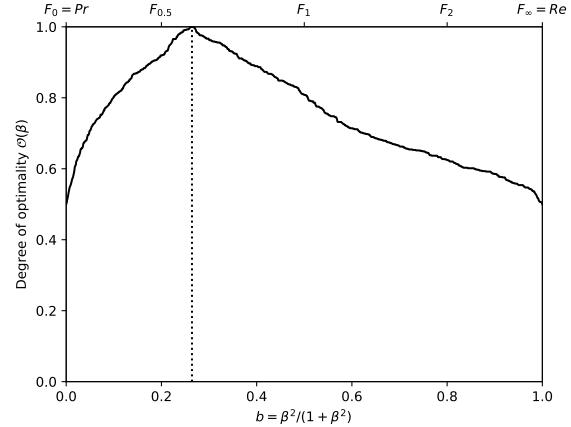
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



(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 .

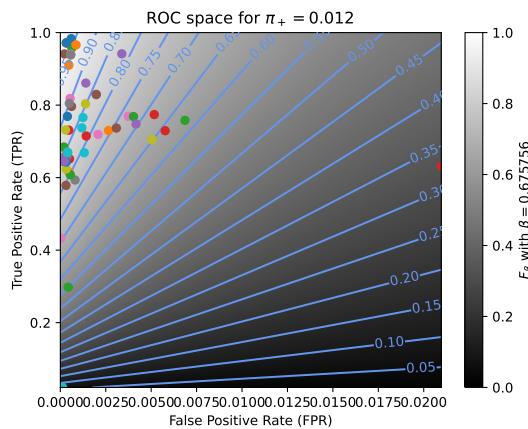


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

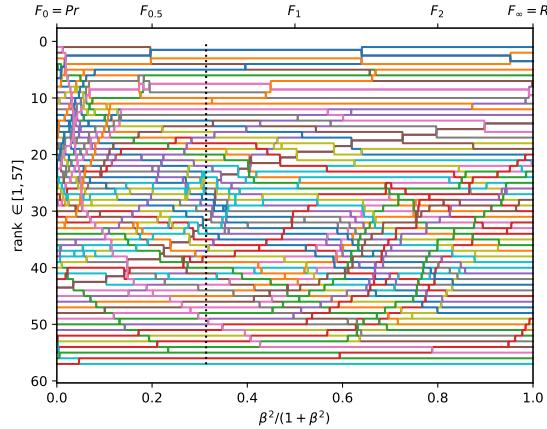


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

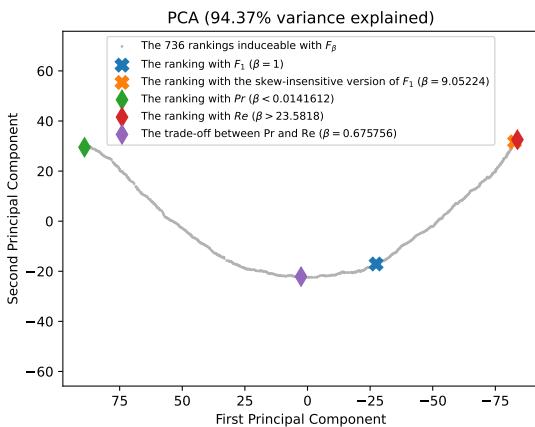
Figure A.3.58. Ranking of 57 BGS methods evaluated on the video "turbulence2" ($\pi_+ = 0.0008$).



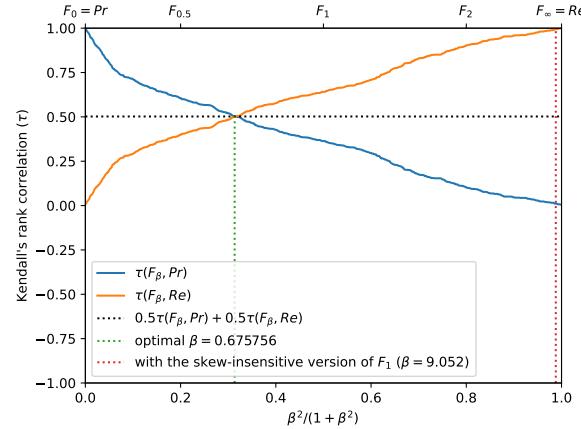
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



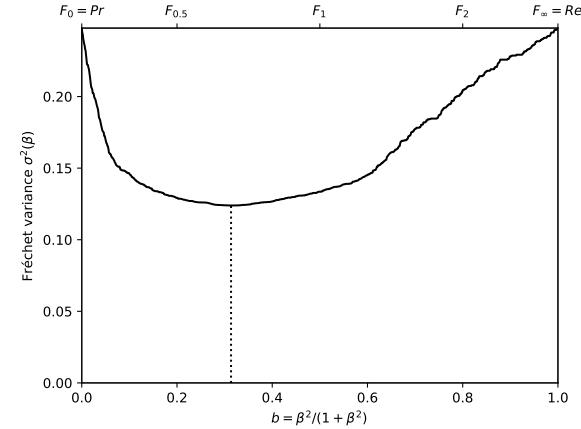
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



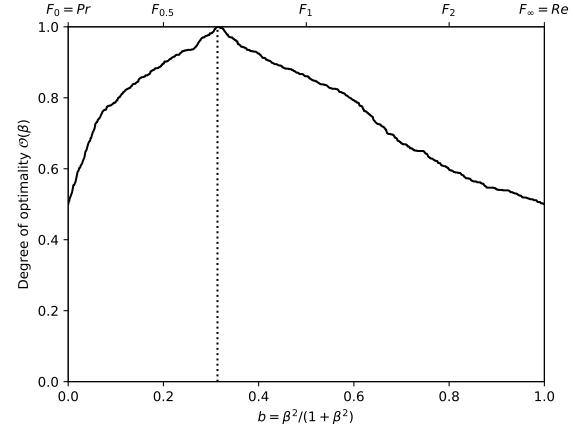
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.

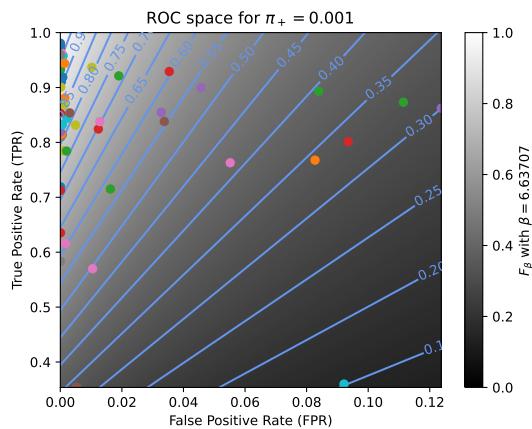


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(F_\beta; Re)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

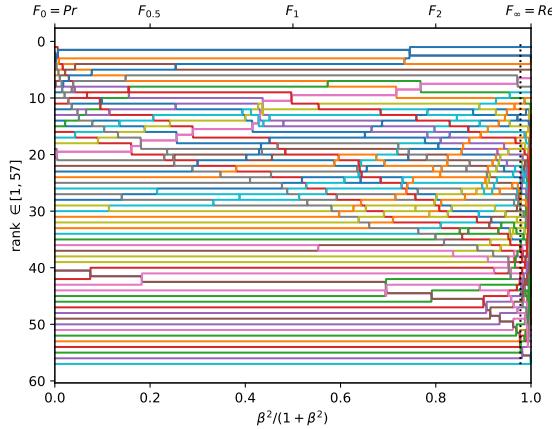


(f) The degree of optimality $O(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (*i.e.*, that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

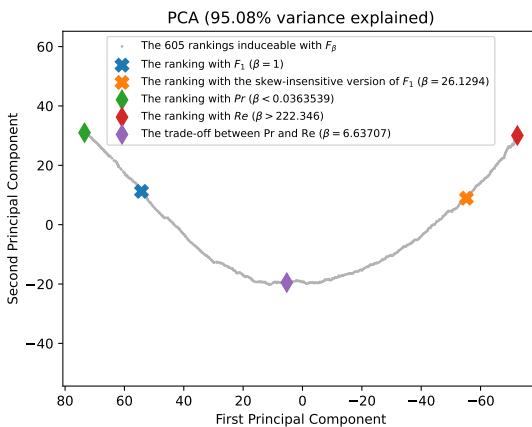
Figure A.3.59. Ranking of 57 BGS methods evaluated on the video "turbulence3" ($\pi_+ = 0.0121$).



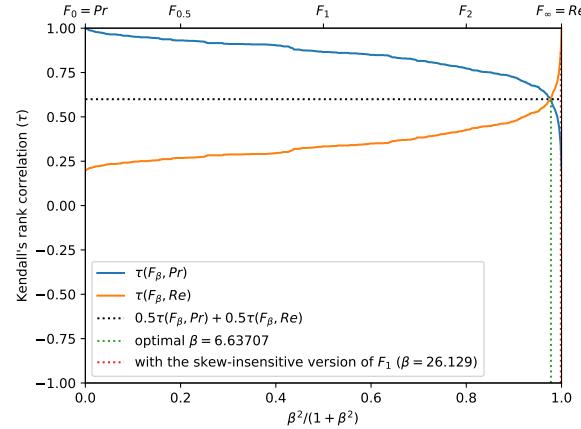
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



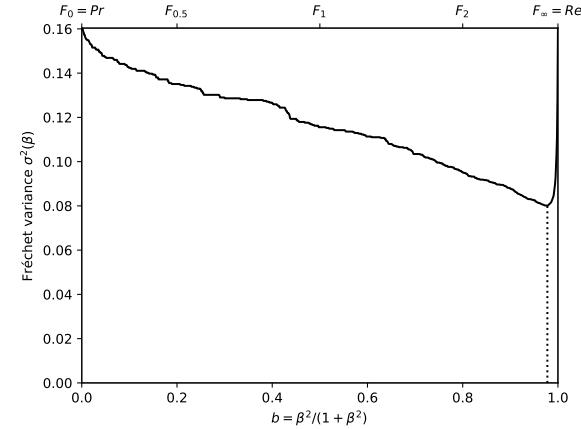
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



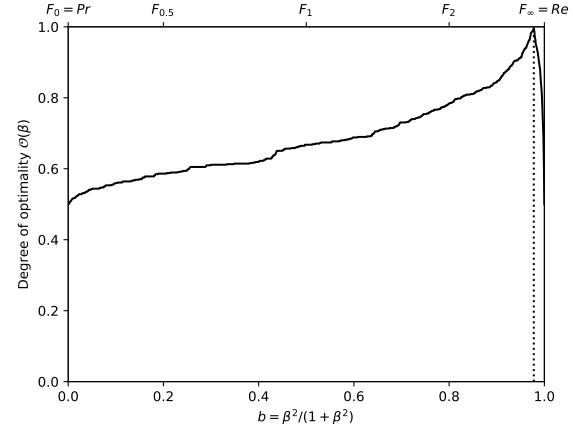
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



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

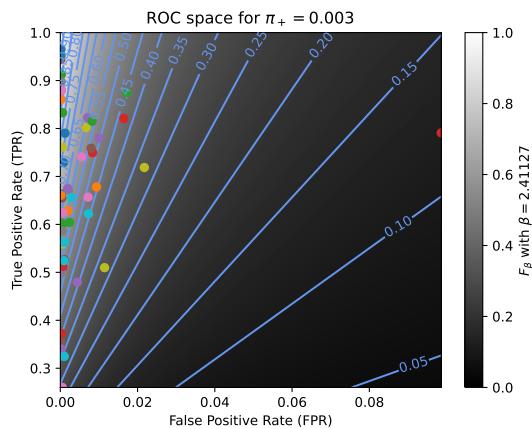


(d) The Fréchet variance $\sigma^2(\beta) = d_\tau^2(Pr; F_\beta) + d_\tau^2(Re; F_\beta)$ w.r.t. β . The optimal value (or range of optimal values) for β is where the curve has its minimum.

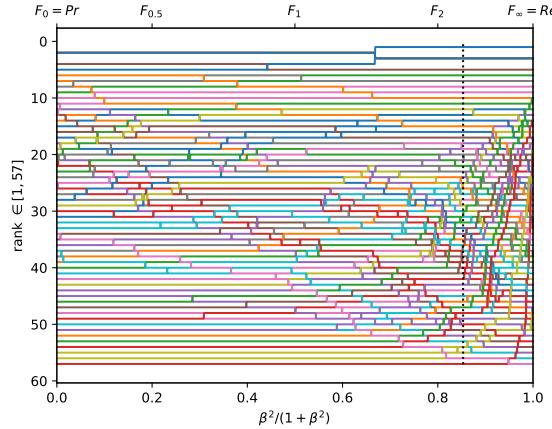


(f) The degree of optimality $\mathcal{O}(\beta)$ w.r.t. β . It is the probability to optimally ordering a pair of classifiers (BGS methods) given that it is not trivial (i.e., that Pr and Re are in contradiction). The optimal value (or range of optimal values) for β is where the curve reaches 100%.

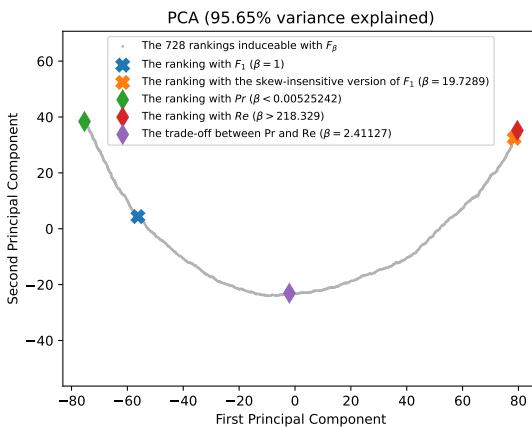
Figure A.3.60. Ranking of 57 BGS methods evaluated on the video "turbulence0" ($\pi_+ = 0.0015$).



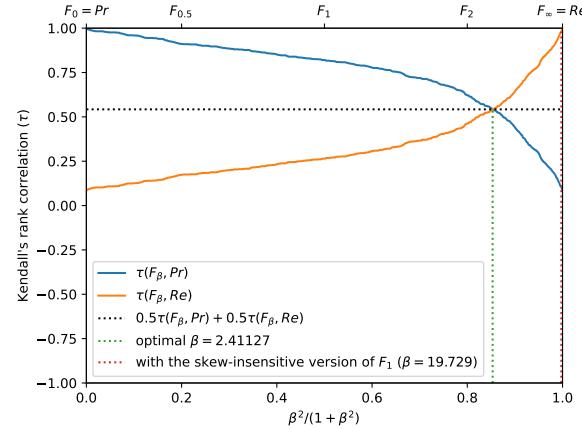
(a) The performances of 57 classifiers (BGS methods) depicted as points in the ROC space, with the isometrics of the optimal tradeoff score, from the ranking point of view, between precision and recall. See Eq. (15).



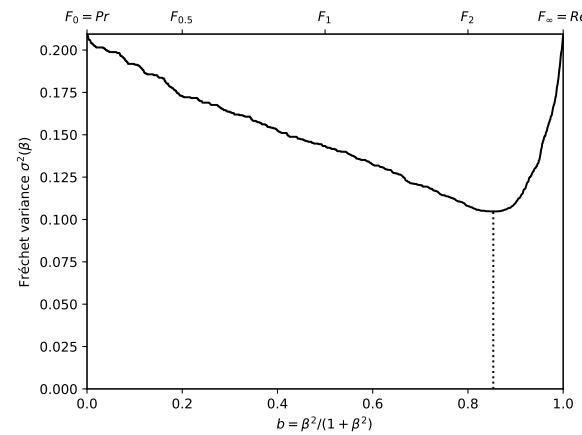
(c) The ranks of each classifier w.r.t. β . The optimal value (or range of optimal values) for β , shown here by the vertical line, is such that the number of swaps on its left is equal to the number of swaps on its right.



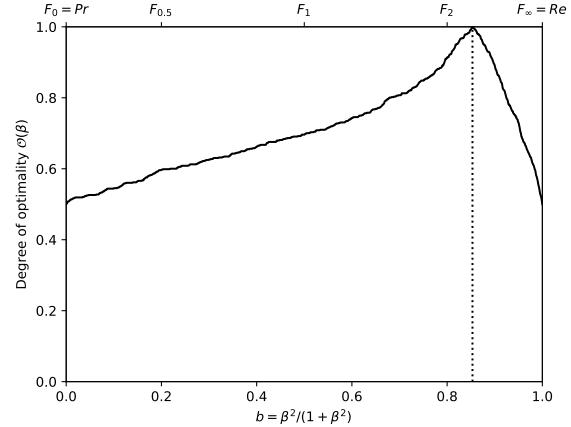
(e) Linear projection (PCA) of the manifold of the rankings induced by the F_β scores. The color points indicate the precision, the recall, F_1 , $SIVF$, as well as the optimal tradeoff. The optimal tradeoff is at the same distance of the two extremities when the distance is measured along the manifold,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



(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.



(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.



(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 "turbulence1" ($\pi_+ = 0.0026$).