Hodnocení výkonnosti klasifikátoru

Nature Inspired
Technologies Group

Nature Inspired
Technologies Group

Matice záměn

	Předvídaný pozitivní ⊕	Předvídaný negativní ⊗	Frekvence podle tříd (rate)
Skutečně pozitivní ⊕	***a***	***b***	$\boldsymbol{TPR} = \boldsymbol{a}/(\boldsymbol{a}+\boldsymbol{b})$ **senzitivita** $FNR = b/(a+b)$
Skutečně negativní ⊗	***c***	***d***	$\boldsymbol{TNR} = \boldsymbol{d}/(\boldsymbol{c}+\boldsymbol{d})$ **specificita** $FPR = c/(c+d)$

$$\textbf{celková\_správnost}\ (accuracy) = (a+d)/(a+b+c+d)$$

$$\boldsymbol{TPR} + \boldsymbol{FNR} = 1,\ \boldsymbol{TNR} + \boldsymbol{FPR} = 1$$

Předv. / skut.	⊕	⊗	Σ
⊕	30	20	**50**
⊗	10	40	**50**
Σ	**40**	**60**	**100**

$\boldsymbol{pos} = (a+b)/(a+b+c+d)$,
$\boldsymbol{neg} = (c+d)/(a+b+c+d)$,

t.j. frekvence skutečně pozitivních/ negativních ve všech příkladech

Předv. / skut.	⊕	⊗	Σ
⊕	30	20	**50**
⊗	10	40	**50**
Σ	**40**	**60**	**100**

$$\textbf{celková\_správnost} = pos * \boldsymbol{TPR} + neg * \boldsymbol{TNR}$$

Příklad permisivního spam filtru,

který nic nepovažuje za spam

Předv. / Skut.	+	X	∑
+	0	10	**10**
X	0	990	**990**
∑	**0**	**1000**	**1000**

Celková_správnost = Accuracy

= (***a*** + ***d***)/(***a***+***b*** +***c*** + ***d***)

99% je skvělá, ale model je k ničemu!

Přesnost (*důvěra, konfidence, precision*) = ***a*** /(***a***+***c***)

„podíl správně klasifikovaných objektů ve všech příkladech zařazených modelem jako ⊕" se používá jako doplňující informace k ***TPR*** v případě velmi nerovnoměrného poměru mezi *Pos* a *Neg*.

V uvedeném příklade ani použití přesnosti nepomůže! Lepší charakteristika je

Úplnost (*recall*) = ***senzitivity*** = ***a*** /(***a***+***b***), zde ***Úplnost*** = 1

Senzitivita a specificita se běžně používají v lékařství:

Senzitivita = ***a*** /(***a***+***b***) „na kolik skutečně nemocných lék zabere"

Specificita = ***d*** /(***c*** +***d***) „jak moc zabírá lék i na zdravé"

Příklad spam filtru

Předv. Skut.	+	X	∑
+	4	6	**10**
X	2	988	**990**
∑	**6**	**994**	**1000**

Accuracy = (***a*** + ***d***)/(***a***+***b*** +***c*** + ***d***)

= 992/1000 = 99,2%

Přesnost (*důvěra, konfidence, precision*)

= ***a*** /(***a***+***c***) … = 4/6 = 66,6 %

tpr = 40 % ***tnr*** = 99,8 %

Který z obou filtrů je lepší?

Předv. Skut.	+	X	∑
+	10	0	**10**
X	8	982	**990**
∑	**18**	**982**	**1000**

Accuracy = 99,2 %

Přesnost = *55,6* %

tpr = 100 % ***tnr*** = 99,2 %

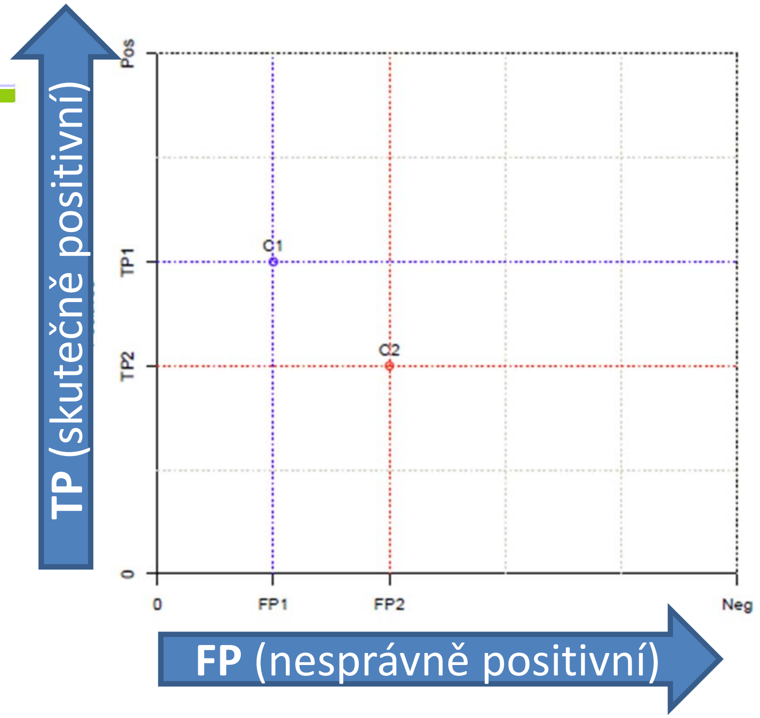
Křivka pokrytí
(*coverage plot*)

Nástroj pro vizualizaci výkonu klasifiká-toru „po klas.třídách" ***TPR*** i ***FPR*** :

Klasifikátor **C1 dominuje** nad **C2**, když **C1** je lepší než **C2** na obou třídách, tj.

$$a/(a+b) = \mathbf{TPR1 > TPR2}$$

a $$c/(c+d) = \mathbf{FPR1 < FPR2}$$



C1

Předv. Skut.	+	X	∑
+	30	20	**50**
X	10	40	**50**
∑	**40**	**60**	**100**

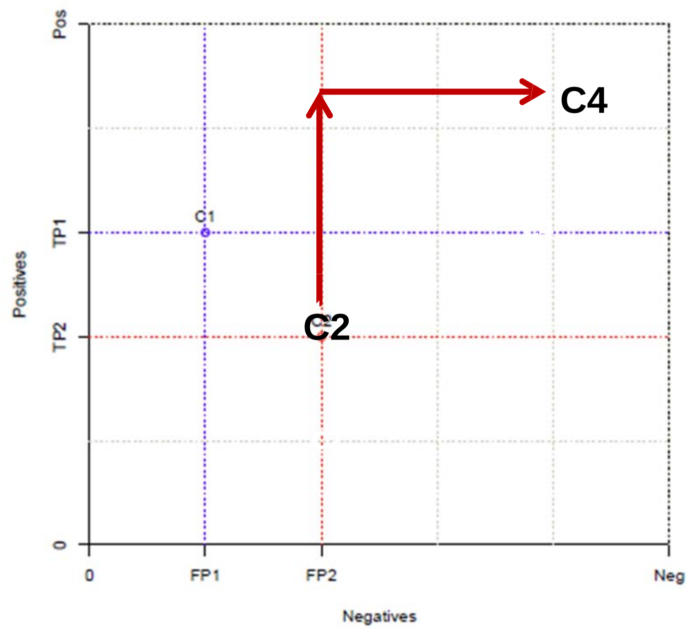
C2

Předv. Skut.	+	X	∑
+	20	30	**50**
X	20	30	**50**
∑	**40**	**60**	**100**

Graf pokrytí

Lze poznat v použitém znázornění, kdy má klasifikátor **C4** pro stejná data stejnou **správnost** (*accuracy*) jako **C2 ?**

$accuracy = (a+d)/(a+b+c+d)$



Je-li výkon **C2** charakterizován tabulkou a),
pak zvýší-li klasifikátor **C4** o *x* počet **True_Positive,** může **C4** dosáhnout stejné správnosti jako **C2** (pro nezměněné hodnoty **Pos,Neg** a **N**) pouze v případě, že počet **TNegative** je o *x* menší a **FP** o *x* větší, viz tab. b)

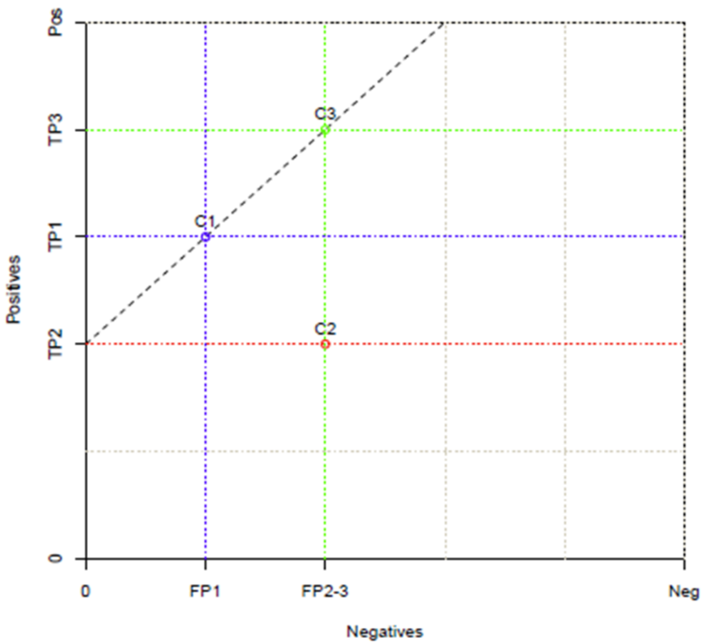
a)

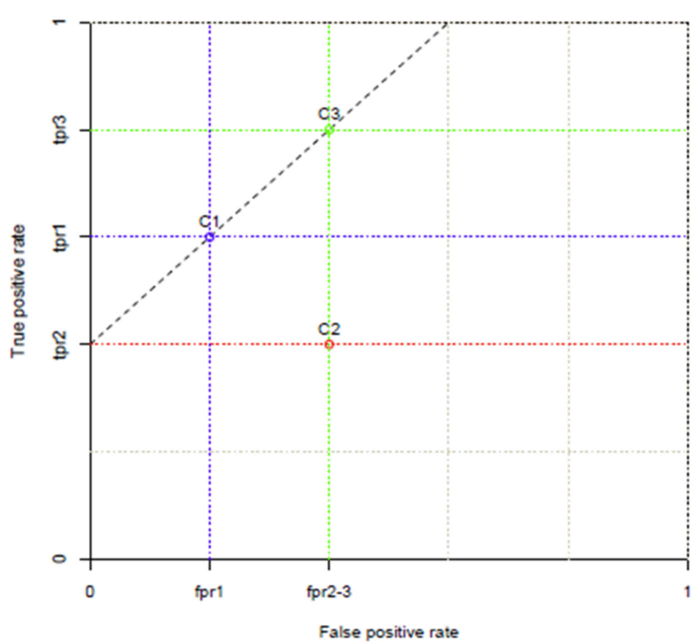
	Předv. Skut.	+	X	Σ
C2	+	*a*	*b*	Pos
	X	*c*	*d*	Neg
	Σ			N

b)

	Předv. Skut.	+	X	Σ
C4	+	*a + x*	*b - x*	Pos
	X	*c + x*	*d - x*	Neg
	Σ			N

Některé vlastnosti graf pokrytí



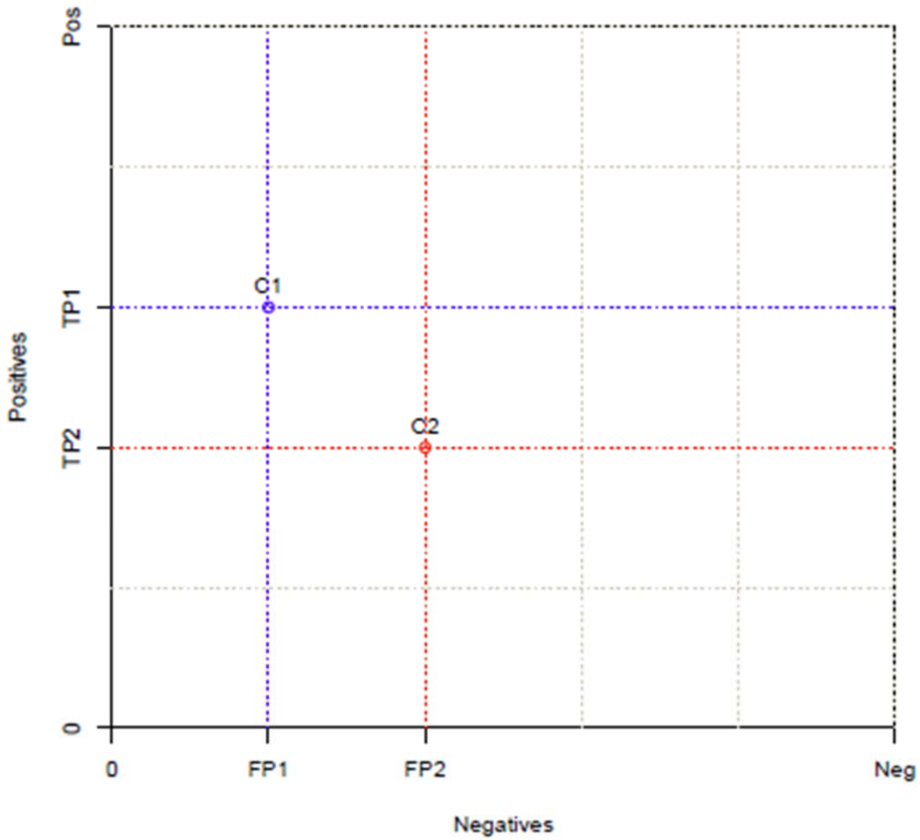


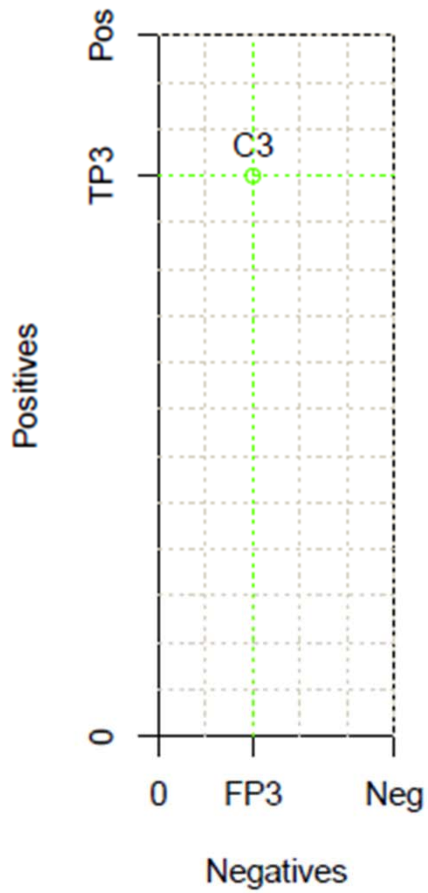
Přesvědčili jsme se, že klasifikátory, jejichž spojnice má na křivce pokrytí směrnici 1, mají stejnou správnost !

Mají-li **C1** a **C3** stejnou správnost, který z nich je lepší?

Je-li důležitější ***TPR***, pak **C3 !**

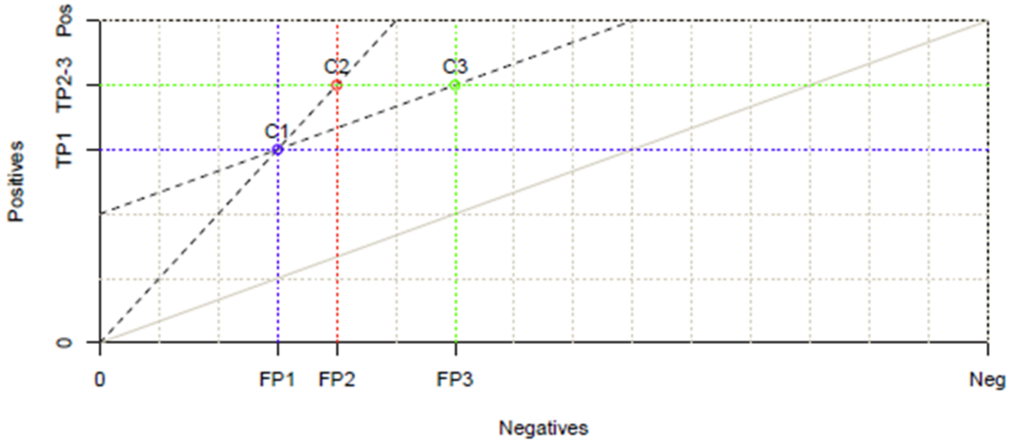
Lze použít graf pokrytí i v případě různé velikosti množin positivních a negativních příkladů ?





Jak srovnávat výkon klasifikátorů nad různými testovacími množinami, ve kterých je nerovnoměrné zastoupení tříd ?

Nerovnoměrné zastoupení tříd



I zde platí, že klasifikátory **C1** a **C2**, jejichž spojnice má na grafu pokrytí směrnici 1, mají **stejnou správnost** !

Co znamená, že spojnice klasifikátorů **C1** a **C3** je na v grafu pokrytí **rovnoběžná s diagonálou**?

Takové **C1** a **C3** mají stejnou **průměrnou přesnost** (average recall) vzhledem k oběma klasifikačním třídám, tj.

$(\boldsymbol{a}/(\boldsymbol{a}+\boldsymbol{b}) + \boldsymbol{d}/(\boldsymbol{c}+\boldsymbol{d}))\ /2$!

Proč je tato hodnota stejná?

Nerovnoměrné zastoupení klasifikačních tříd a graf pokrytí

Předpokládejme, že chceme srovnat výkon 2 klasifikátorů **C1** a **C3** na jediné skupině klasifikovaných dat monutnosti **N**, ve které je **Pos** positivně klasifikovaných příkladů a **Neg** negativně a platí, že **Pos** a **Neg** jsou **různé** a platí **Pos + Neg = N.** Je-li výkon **C1** charakterizován tabulkou a), pak zvýší-li klasifikátor **C4** o ***x*** počet **TP** oproti **C1,** může součet **Pos** být zachován jen v případě, že **FN** bude o ***x*** nižší, viz tabulka b). Podobně v zroste-li **FP** o ***y***, musí ***TN*** o ***y*** klesnout, aby zůstala zachována hodnota **Neg**.

a) C1

Předv. / Skut.	+	X	∑
+	*A*	*b*	Pos
X	*c*	*d*	Neg
∑			N

b) C3

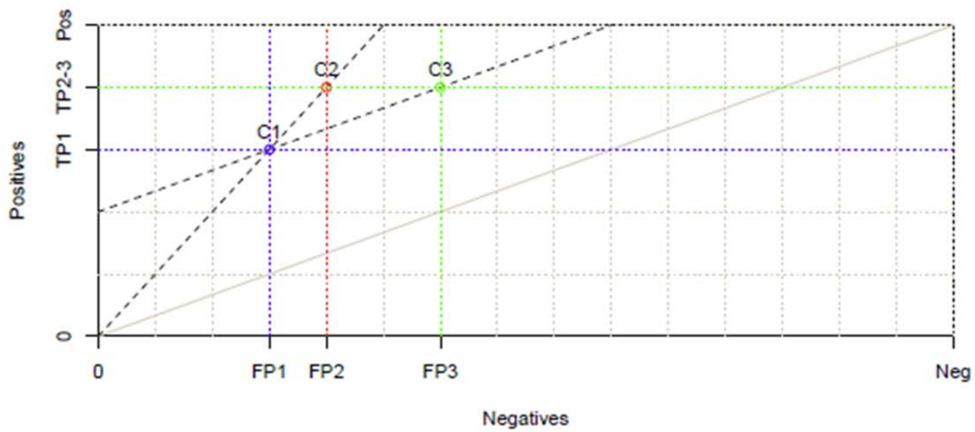
Předv. / Skut.	+	X	∑
+	$a + x$	$b - x$	Pos
X	$c + y$	$d - y$	Neg
∑			N

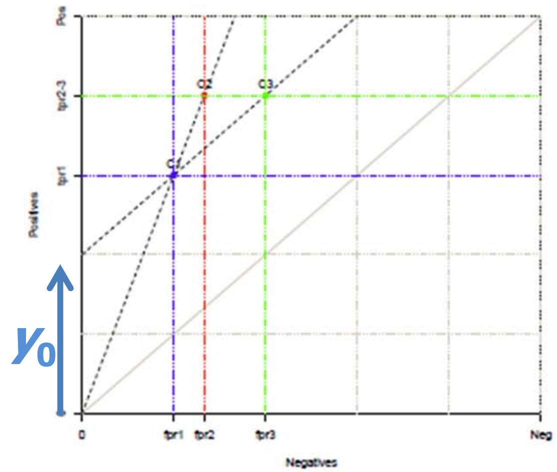
Průměrná_přesnost(**C3**) = $0{,}5 * ((a+x)/(a+b) + (d-y)/(c+d)) =$
$= 0{,}5*((a/(a+b)+d/(c+d)) + (x/(a+b)-y/(c+d))=$**P_p**(**C1**) $+ 0{,}5* (x/(a+b)-y/(c+d))$

Pokud $x/(a+b)-y/(c+d) = 0$, pak **Prům_přesnost**(**C3**)= **Prům_přesnost**(**C1**). To nastává právě když $x / y = (a+b)/c+d) =$ **Pos/Neg**, tedy spojnice obrazů **C1** a **C3** je rovnoběžná s diagonálou v grafu pokrytí.

Znázornění při nerovnoměrném zastoupení tříd?

Vhodným řešením je **normalizace**, tj. hodnoty na osách budou ***TPR*** a ***FPR*,** místo ***TP*** a ***FP*.** Výsledkem je **ROC zobrazení.**





U **ROC** platí, že klasifikátory, jejichž **spojnice má směrnici**

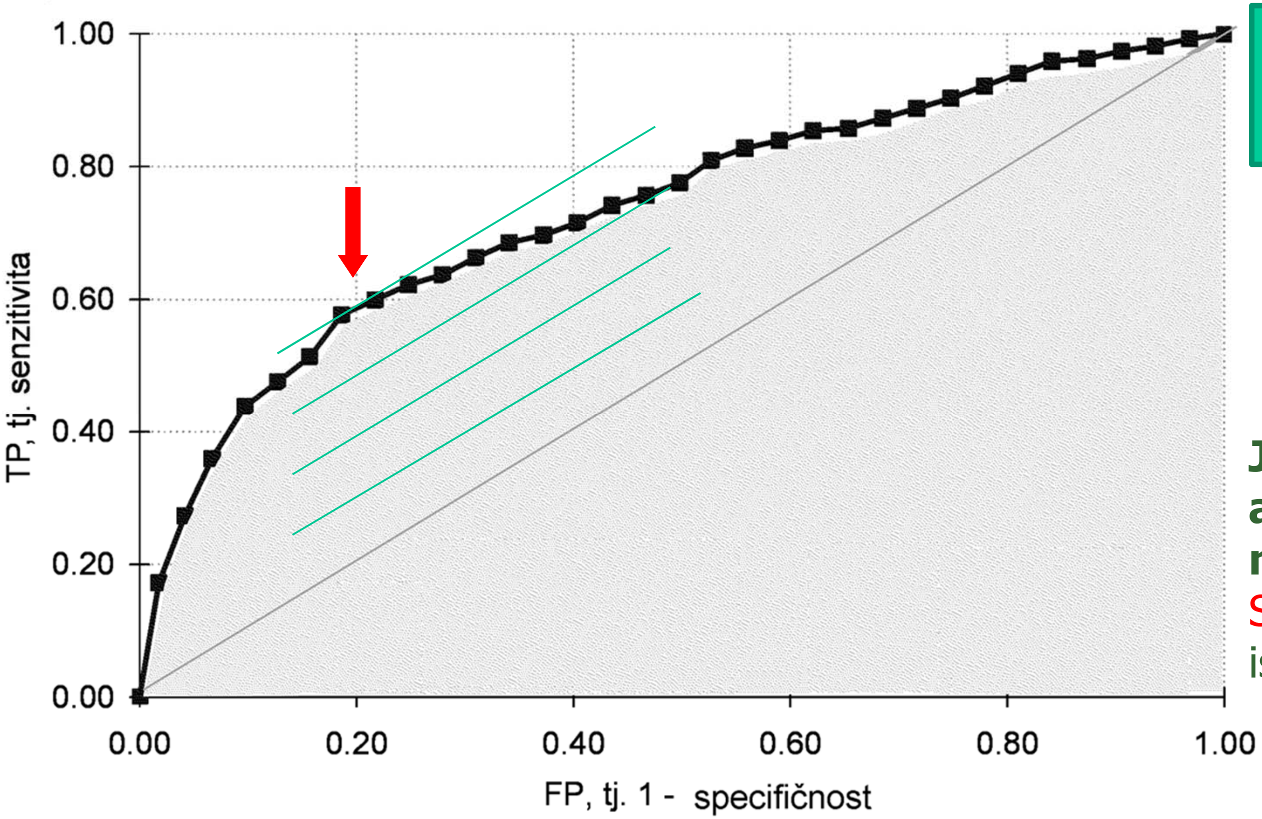
- ***Neg/Pos*** , mají stejnou **správnost (celkovou přesnost)**
- **1**, mají stejnou **průměrnou přesnost:** Pro libovolný bod na takové přímce platí $tpr = y_0 + fpr = y_0 + (1 - tnr)$, z čehož plyne $tpr + tnr = (y_0 + 1)$. Z toho vyplývá, že **průměrná přesnost** je stejná pro všechny body přímky.

Takovým přímkám se říká **isomery**

Nature Inspired Technologies Group

Křivka ROC (Receiver Operating Char.)

	Předvídaný pozitivní	Předvídaný negativní	Frekvence podle tříd (rate)
Skutečně pozitivní	***a***	***b***	***TPR*** = ***a***/(***a*** + ***b***) **senzitivita** *FNR* = *b*/(*a* + *b*)
Skutečně negativní	***c***	***d***	***TNR*** = ***d***/(***c*** + ***d***) **specificita** *FPR* = *c*/(*c* + *d*)



správnost = (a+d)/(a+b+c+d)

- Vypočtené výkony klasifikátorů
- Isoméry rovnoběžné s přímkou mezi body (0,0) a (1,1).

Jak vybrat parametr tak, aby klasifikátor dosahoval nejvyšší možné správnosti?

Správný výběr prostřednictvím isomér (viz předchozí snímek).

Křivka ROC (Receiver Operating Char.)

- Srovnání modelů neprovádíme v jediném bodě
- **Vhodné u modelů, které přímo neklasifikují, ale odhadují pravděpodobnost příslušnosti k některé ze tříd.**

- Např. mějme klasifikační algoritmus, který predikuje pravděpodobnost příslušnosti k pozitivní třídě $p(\boldsymbol{x}(i))$. O diskrétní klasifikaci rozhoduje parametr θ pro hodnotu prahu:
  - Je-li pro objekt i hodnota $p(\boldsymbol{x}(i)) > \theta$, zařadíme objekt popsaný vektorem nezávislých veličin $\boldsymbol{x}(i)$ mezi pozitivní příklady,
  - je-li menší, označíme jej jako negativní.

Pro různé hodnoty θ vzniknou různé klasifikátory. Jak volit hodnotu θ tak, aby správnost algoritmu byla co nejvyšší?

Příklad konstrukce ROC křivky

Pro $\theta = 0{,}22$ je

TP=4, tedy TPR=0,8

FP=2, tedy FPR=0,4

Pro $\theta = 0{,}3$ je

TP=3, tedy TPR=0,6

FP=2, tedy FPR=0,4

Pro $\theta = 0{,}31$ je

TP=3, tedy TPR=0,6

FP=1, tedy FPR=0,2

Pro $\theta = 0{,}4$ je

TP=3, tedy TPR=0,6

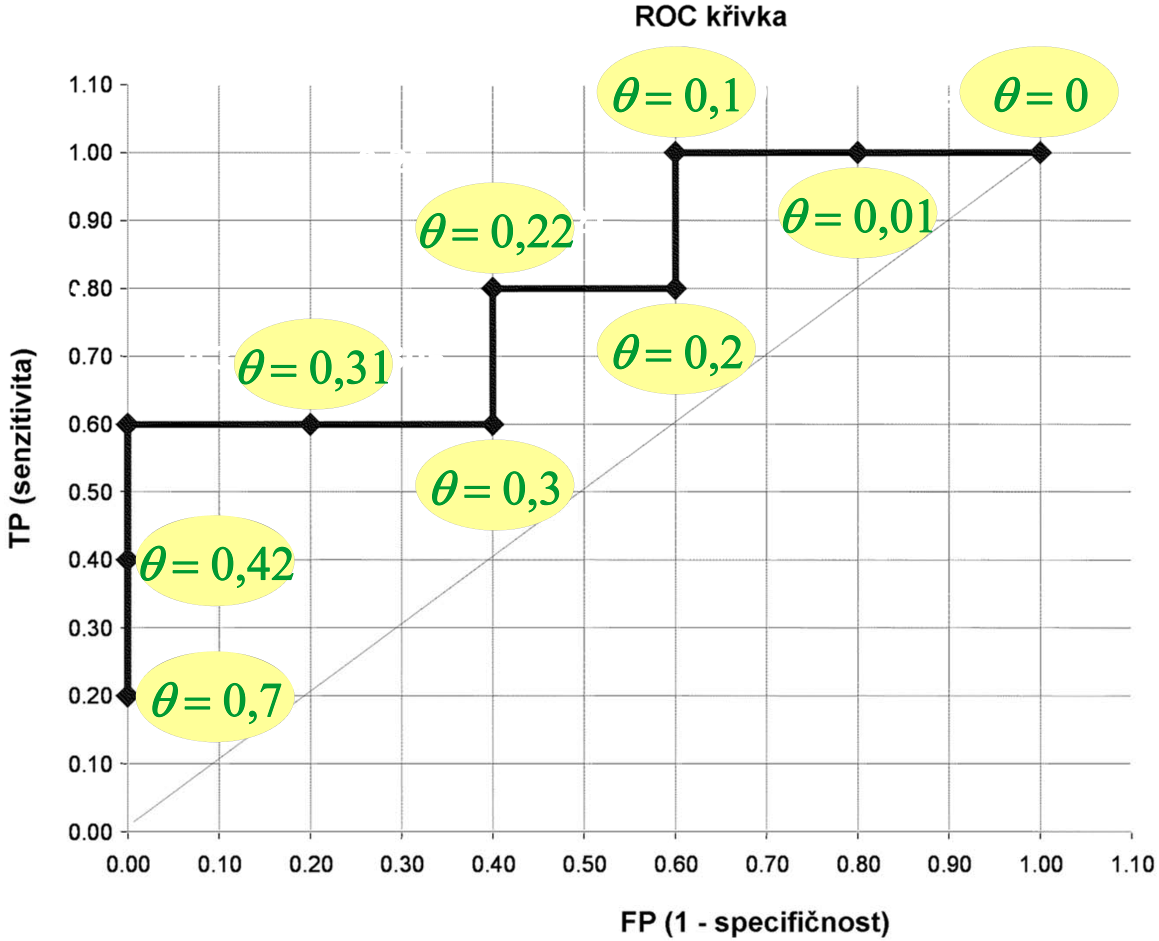
FP=0, tedy FPR=0

objekt	1	2	3	4	5	6	7	8	9	10
predikce	0	0,05	0,2	0,22	0,3	0,31	0,4	0,42	0,7	0,8
skutečnost	0	0	1	0	1	0	0	1	1	1

Pozit. klas. pro $\theta = 0{,}22$

Křivka ROC vznikne vypočtením hodnot *TP* a *FP* pro **všechny různé prahy** θ: 0, 0,01; 0,1; ..; 0,75

Užitečnost různých hodnotících funkcí (dávajících predikci) lze srovnávat podle jejich parametru **AUC** = „plocha pod křivkou".



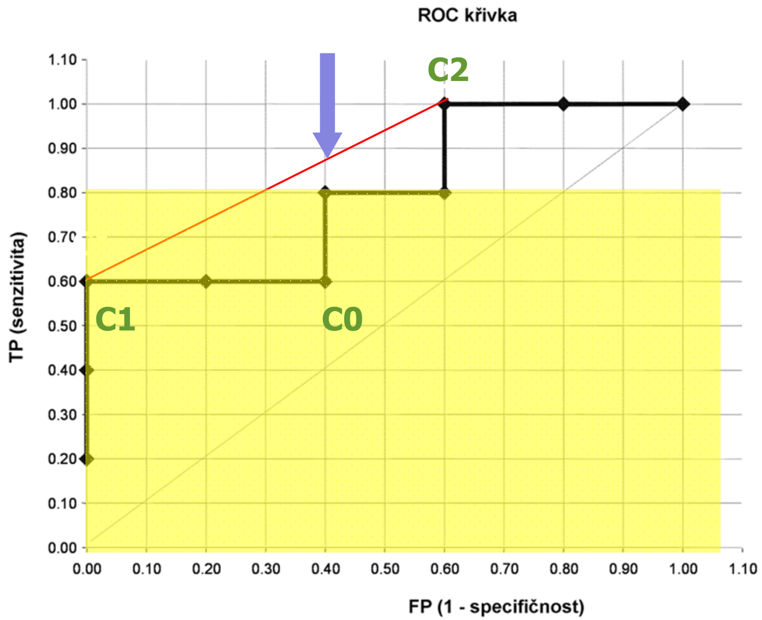
Může ROC křivka pomoci při konstrukci lepších klasifikátorů?

Existuje klasifikátor, který dominuje klasifikátoru **C0** a má přitom *TPR* nejméně 0,8?

Ano, lze jej zkonstruovat jako klasifikátor kombinující rozhodnutí **C1** s pravděpodobností α a **C2** s pravděpodobností (1- α):

- pro α **= 1** je výsledný klasifikátor totožný s **C1**,
- pro α **= 0** zase **C2.**

Tato kombinace dává všechny možnosti na spojnici bodů **C1** a **C2** – můžeme tedy pracovat s **konvexním obalem** původní ROC křivky, na kterém nalezneme vhodnou hodnotu.



Doporučení pro volbu α :

$\alpha = (FP2-FP0)/(FP2-FP1)$