



B4B33RPH: Řešení problémů a hry

Spam filtr - vyhodnocení

Petr Pošík

Katedra kybernetiky

ČVUT FEL



Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti



Abychom si rozuměli...

Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

- Abychom si rozuměli...
- Hodnocení klasifikátoru
- Vizualizace úspěšnosti
- Kvíz

Spam filtr

Shrnutí

Dotazy?

Zadání: máme

- datovou sadu s jistým známým počtem
 - pozitivních případů (např. spamů) a
 - negativních případů (např. hamů), a
- klasifikátor (spam filtr), který pro každý případ v datové sadě poskytne svůj odhad (předpověď, predikci), zda se jedná o pozitivní či negativní případ.

Chceme zjistit úspěšnost klasifikátoru na dané datové sadě.

	V datové sadě	
	pozitivních	negativních
Pozitivní předpověď	TP	FP
Negativní předpověď	FN	TN
Celkem	$TP + FN$	$FP + TN$

True positives (TP): počet případů klasifikátorem *správně* označených jako *pozitivní*.

False positives (FP): počet případů klasifikátorem *chybně* označených jako *pozitivní*.

False negatives (FN): počet případů klasifikátorem *chybně* označených jako *negativní*.

True negatives (TN): počet případů klasifikátorem *správně* označených jako *negativní*.



Hodnocení klasifikátoru

Přesnost: poměr správných klasifikací ku celkovému počtu případů

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \quad (1)$$

Jak měřit zvlášť úspěšnost při klasifikaci pozitivních a negativních případů?

Můžeme použít např. dvojici *TP rate* a *FP rate*:

$$TP\ rate = \frac{TP}{TP + FN} = 1 - FN\ rate \quad (2)$$

$$FP\ rate = \frac{FP}{TN + FP} = 1 - TN\ rate \quad (3)$$

Naše míra kvality q (upravená přesnost Acc):

$$q = \frac{TP + TN}{TP + FN + TN + 10 \cdot FP} \quad (4)$$

Lze tyto tři míry (*TP rate*, *FP rate*, q) nějakým způsobem jednoduše vizualizovat?

Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

- Abychom si rozuměli...
- **Hodnocení klasifikátoru**
- Vizualizace úspěšnosti
- Kvíz

Spam filtr

Shrnutí

Dotazy?



Vizualizace úspěšnosti

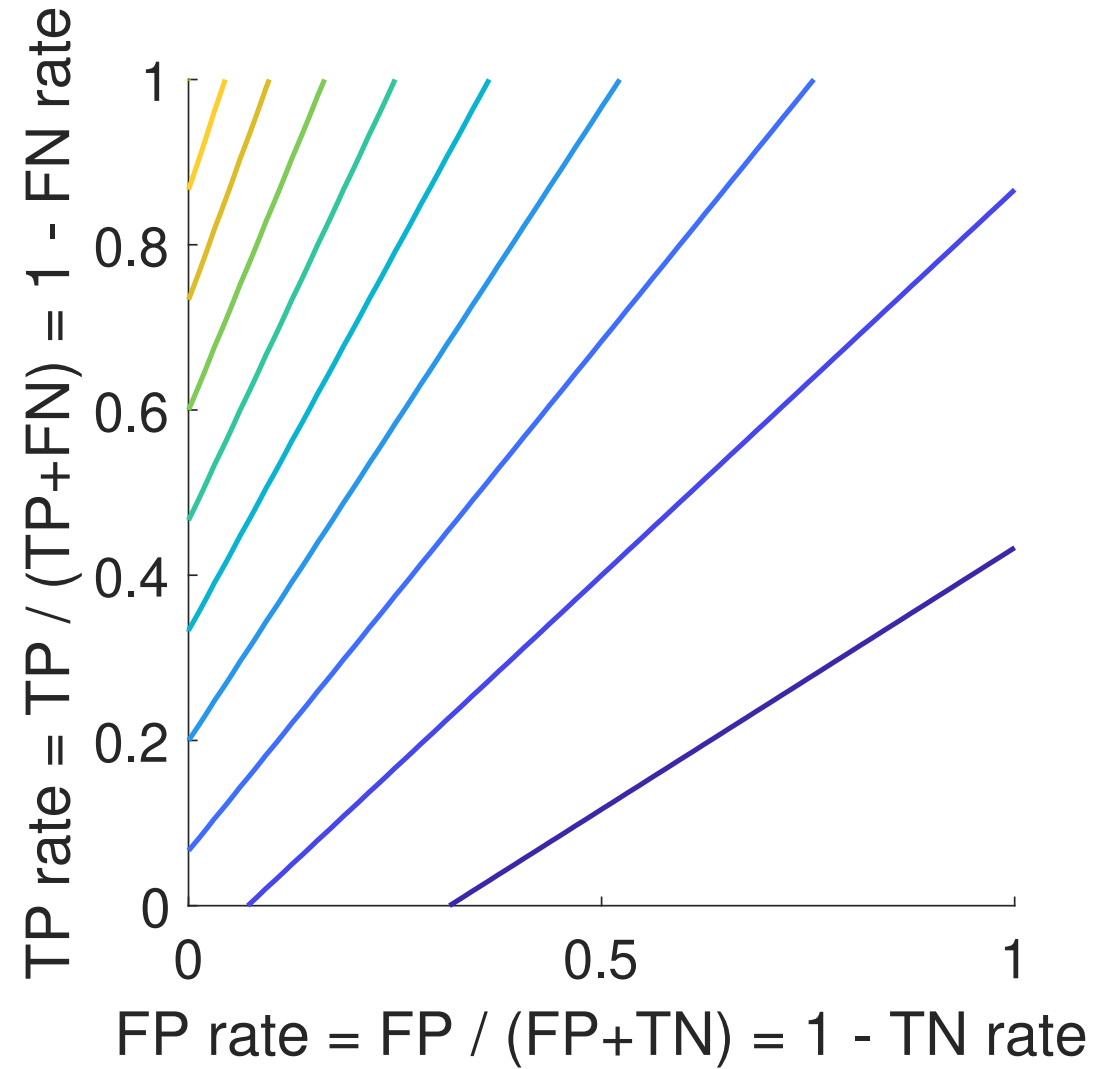
Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

- Abychom si rozuměli...
- Hodnocení klasifikátoru
- **Vizualizace úspěšnosti**
- Kvíz

Spam filtr

Shrnutí

Dotazy?





Kvíz

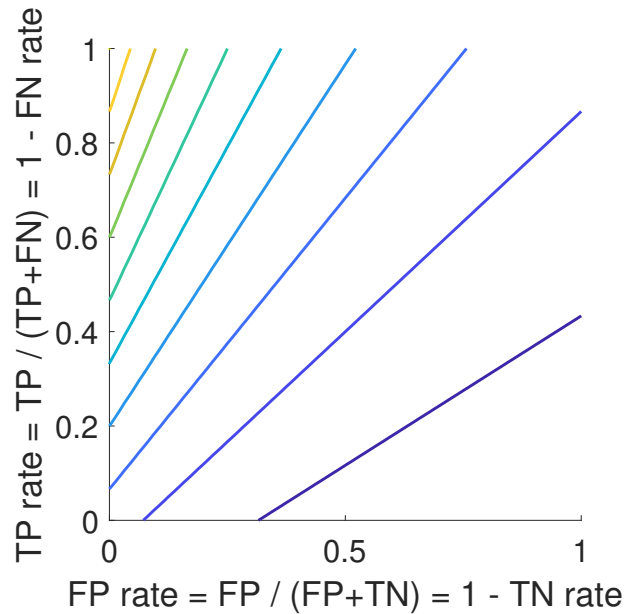
Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

- Abychom si rozuměli...
- Hodnocení klasifikátoru
- Vizualizace úspěšnosti
- **Kvíz**

Spam filtr

Shrnutí

Dotazy?



Kde v tomto grafu bude ležet bod znázorňující "Paranoidní" spam filtr, tedy takový, který všechny emaily zařadí do pozitivní třídy (spam)?

- A** V bodě (0, 0), vlevo dole.
- B** V bodě (1, 0), vpravo dole.
- C** V bodě (0, 1), vlevo nahoře.
- D** V bodě (1, 1), vpravo nahoře.



Spam filtr



Trocha suché statistiky

Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

Spam filtr

- Trocha suché statistiky

- 3 datové sady
- Bodové hodnocení
- Data 1: výsledky
- Data 2: výsledky
- Data 3: výsledky
- Porovnání výsledků na 3 sadách dat
- Vítězné filtry

Shrnutí

Dotazy?

Pozice úlohy “Spam filtr” v předmětu RPH:

- 2009: 2 úlohy / semestr, spam filtr volitelný ze 3 úloh
- 2010: 3 úlohy / semestr, spam filtr volitelný ze 2 úloh
- 2011-nyní: 3 úlohy / semestr, spam filtr povinný
- 2020: Spam filtr - týmová úloha

Celkem: celkový počet studentů zapsaných na RPH

Odevzdáno: počet studentů, kteří odevzdali spam filtr

Bez chyb: počet studentů, jejichž spam filtry fungovaly bez chyby

Rok	Celkem		Odevzdáno		Bez chyb	
2009	157	100.0 %	16	10.2 %	15	9.6 %
2010	90	100.0 %	21	23.3 %	21	23.3 %
2011	118	100.0 %	104	88.1 %	102	86.4 %
2012	110	100.0 %	89	80.9 %	87	79.1 %
2013	112	100.0 %	92	82.1 %	89	79.5 %
2014	103	100.0 %	83	80.6 %	81	78.6 %
2015	142	100.0 %	115	80.1 %	112	78.9 %
2016	204	100.0 %	165	80.9 %	155	76.0 %
2017	195	100.0 %	151	77.4 %	138	70.7 %
2018	205	100.0 %	176	85.9 %	167	81.5 %
2019	201	100.0 %	161	80.0 %	153	76.1 %
2020	199	100.0 %	164	82.4 %	156	78.4 %



3 datové sady

Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

Spam filtr

- Trocha suché statistiky
- **3 datové sady**
- Bodové hodnocení
- Data 1: výsledky
- Data 2: výsledky
- Data 3: výsledky
- Porovnání výsledků na 3 sadách dat
- Vítězné filtry

Shrnutí

Dotazy?

První datová sada (shodná s online automatickým skriptem):

	HAM	SPAM	Celkem	Poměr HAM:SPAM
Trénovací sada	153	461	614	1 : 3
Testovací sada	153	461	614	1 : 3

Druhá datová sada (z jiného zdroje, ale se stejnou velikostí a složením jako první sada):

	HAM	SPAM	Celkem	Poměr HAM:SPAM
Trénovací sada	150	450	600	1 : 3
Testovací sada	150	450	600	1 : 3

Třetí datová sada (stejný zdroj jako 2. sada, větší trénovací sada, HAM i SPAM zastoupen rovnoměrně):

	HAM	SPAM	Celkem	Poměr HAM:SPAM
Trénovací sada	1500	1500	3000	1 : 1
Testovací sada	200	600	800	1 : 3

Datové sady 2 a 3 od roku 2016/17 jiné než předchozí roky!



Bodové hodnocení

Nové datové sady 2 a 3 se ukázaly být pro klasifikaci objektivně těžší než staré datové sady.

■ Sada 1:

Kvalita q	4	0.9	3	0.7	2	0.5	1	0.3	0
Body	4		3		2		1		0

■ Sady 2 a 3:

Kvalita q	4	0.8	3	0.6	2	0.4	1	0.3	0
Body	4		3		2		1		0

Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

Spam filtr

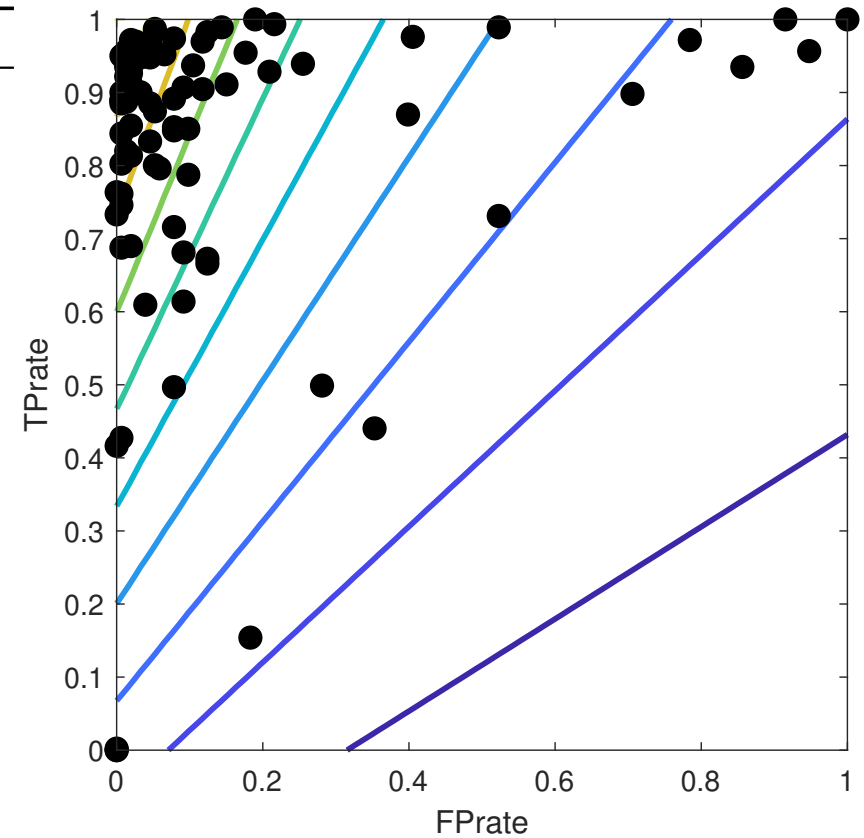
- Trocha suché statistiky
- 3 datové sady
- **Bodové hodnocení**
- Data 1: výsledky
- Data 2: výsledky
- Data 3: výsledky
- Porovnání výsledků na 3 sadách dat
- Vítězné filtry

Shrnutí

Dotazy?

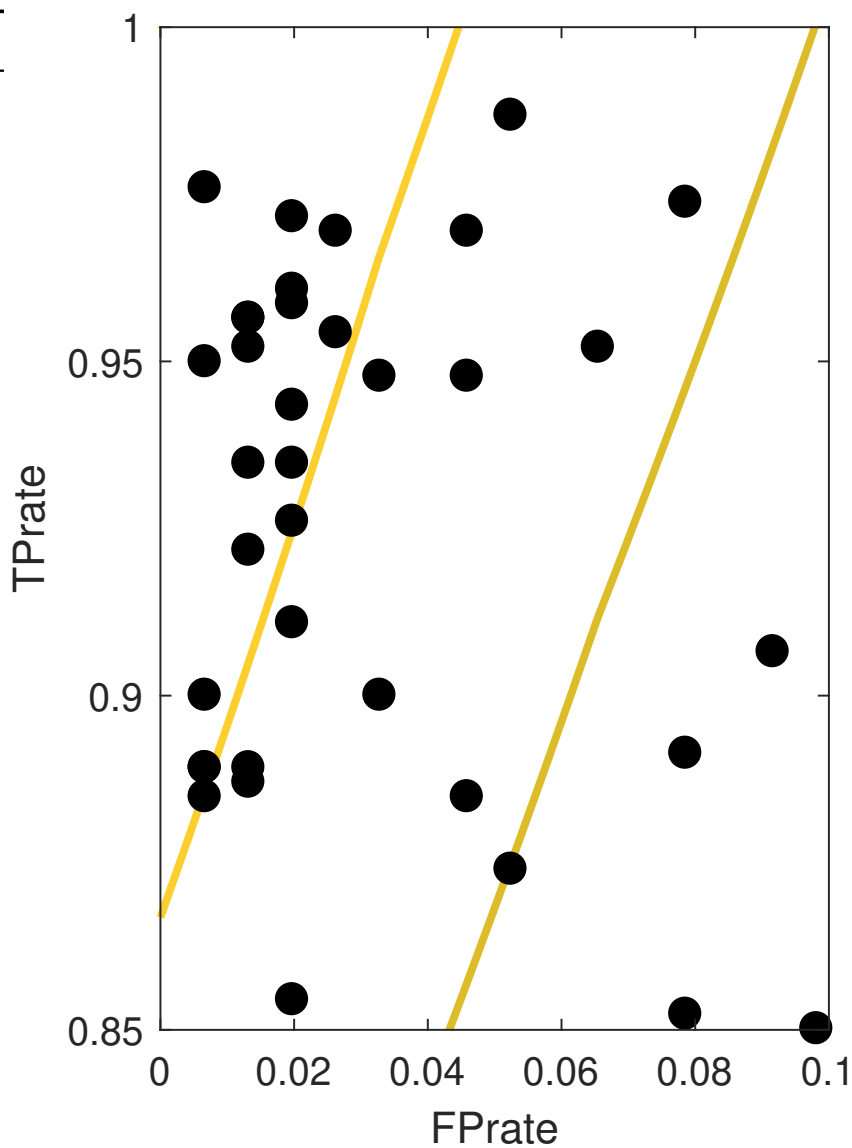
Data 1: výsledky

#	Username	q	FP	FN	FP rate	TP rate
6.0	pajtaand_sobotma6	0.947	1	23	0.007	0.950
7.5	cezneluk_libalrud	0.937	2	20	0.013	0.957
7.5	blahoja9_bodnafab	0.937	2	20	0.013	0.957
9.0	nejedmat_vonasmic	0.934	2	22	0.013	0.952
10.0	banovvik_belouste	0.933	3	13	0.020	0.972
11.0	mareska9_novacvi1	0.925	3	18	0.020	0.961
12.0	saranego	0.924	3	19	0.020	0.959
13.0	weisjose	0.921	2	30	0.013	0.935
14.0	lososjos_stanope3	0.917	4	14	0.026	0.970
15.0	valesja4_zderama1	0.913	3	26	0.020	0.944
16.0	vavermic_ziltjura	0.911	2	36	0.013	0.922
17.0	pelikma3	0.910	1	46	0.007	0.900
18.0	korlavik_korladia	0.906	3	30	0.020	0.935
19.0	ospananu	0.906	4	21	0.026	0.954
20.5	nurkharl_poncakar	0.902	1	51	0.007	0.889
20.5	levitvas_nazarboh	0.902	1	51	0.007	0.889
22.0	shubevik	0.900	3	34	0.020	0.926
23.0	urbanst6	0.899	1	53	0.007	0.885
24.0	yazykvla	0.889	3	41	0.020	0.911
25.0	kantsmyk_kozlevla	0.888	5	24	0.033	0.948
26.0	barabda1_uherjir3	0.888	2	51	0.013	0.889
27.0	kaslonat_msalhoma	0.886	2	52	0.013	0.887
28.0	horakj31_milyunik	0.876	7	14	0.046	0.970
29.0	divisvi1_douchmar	0.875	8	6	0.052	0.987
30.0	litvidmi_mamaeand	0.868	1	72	0.007	0.844
...



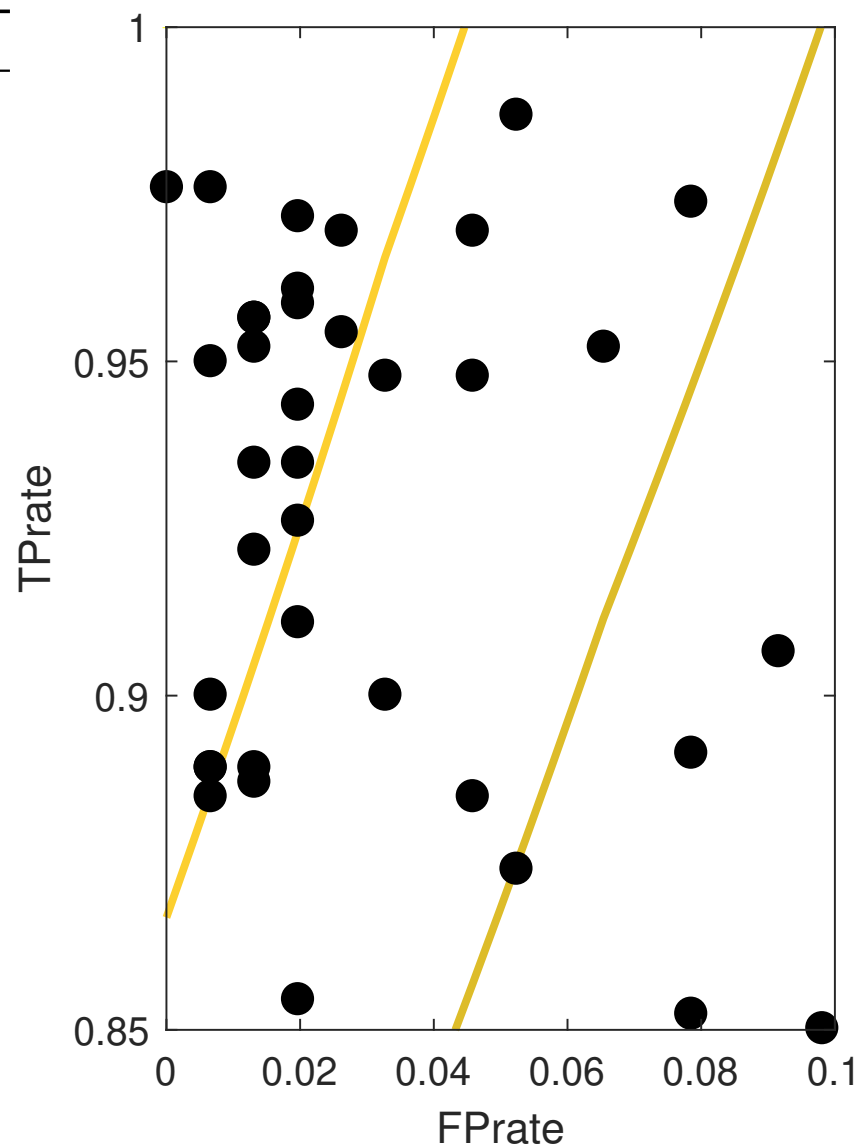
Data 1: výsledky

#	Username	q	FP	FN	FP rate	TP rate
5.0	kovalda2_zribktad	0.966	1	11	0.007	0.976
6.0	pajtaand_sobotma6	0.947	1	23	0.007	0.950
7.5	cezneluk_libalrud	0.937	2	20	0.013	0.957
7.5	blahoja9_bodnafab	0.937	2	20	0.013	0.957
9.0	nejedmat_vonasmic	0.934	2	22	0.013	0.952
10.0	banovvik_belouste	0.933	3	13	0.020	0.972
11.0	mareska9_novacvil	0.925	3	18	0.020	0.961
12.0	saranego	0.924	3	19	0.020	0.959
13.0	weisjose	0.921	2	30	0.013	0.935
14.0	lososjos_stanope3	0.917	4	14	0.026	0.970
15.0	valesja4_zderama1	0.913	3	26	0.020	0.944
16.0	vavermic_ziltjura	0.911	2	36	0.013	0.922
17.0	pelikma3	0.910	1	46	0.007	0.900
18.0	korlavik_korladia	0.906	3	30	0.020	0.935
19.0	ospananu	0.906	4	21	0.026	0.954
20.5	nurkharl_poncakar	0.902	1	51	0.007	0.889
20.5	levitvas_nazarboh	0.902	1	51	0.007	0.889
22.0	shubevik	0.900	3	34	0.020	0.926
23.0	urbanst6	0.899	1	53	0.007	0.885
24.0	yazykvla	0.889	3	41	0.020	0.911
25.0	kantsmyk_kozlevla	0.888	5	24	0.033	0.948
26.0	barabda1_uherjir3	0.888	2	51	0.013	0.889
27.0	kaslonat_msalhoma	0.886	2	52	0.013	0.887
28.0	horakj31_milyunik	0.876	7	14	0.046	0.970
29.0	divisvi1_douchmar	0.875	8	6	0.052	0.987
30.0	litvidmi_mamaeand	0.868	1	72	0.007	0.844
...



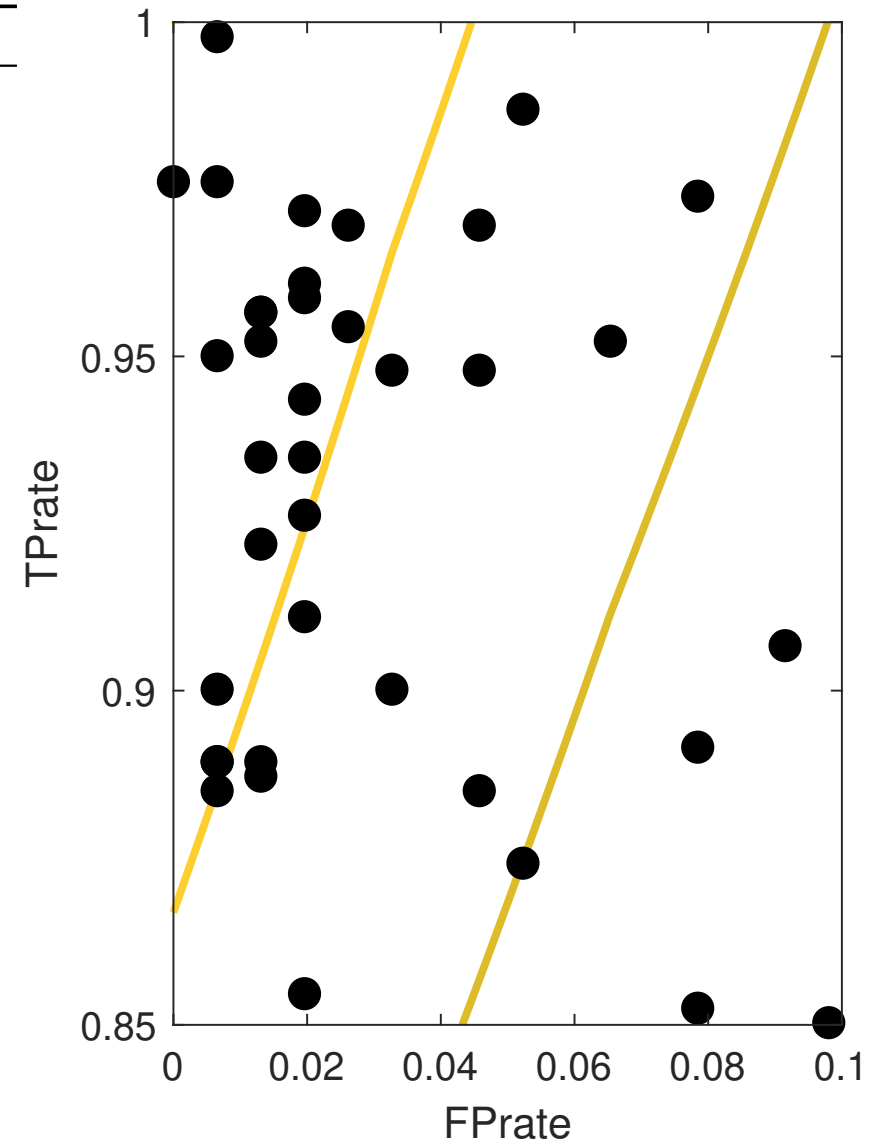
Data 1: výsledky

#	Username	q	FP	FN	FP rate	TP rate
4.0	knoblvit_louckada	0.982	0	11	0.000	0.976
5.0	kovalda2_zribktad	0.966	1	11	0.007	0.976
6.0	pajtaand_sobotma6	0.947	1	23	0.007	0.950
7.5	cezneluk_libalrud	0.937	2	20	0.013	0.957
7.5	blahoja9_bodnafab	0.937	2	20	0.013	0.957
9.0	nejedmat_vonasmic	0.934	2	22	0.013	0.952
10.0	banovvik_belouste	0.933	3	13	0.020	0.972
11.0	mareska9_novacvil	0.925	3	18	0.020	0.961
12.0	saranego	0.924	3	19	0.020	0.959
13.0	weisjose	0.921	2	30	0.013	0.935
14.0	lososjos_stanope3	0.917	4	14	0.026	0.970
15.0	valesja4_zderama1	0.913	3	26	0.020	0.944
16.0	vavermic_ziltjura	0.911	2	36	0.013	0.922
17.0	pelikma3	0.910	1	46	0.007	0.900
18.0	korlavik_korladia	0.906	3	30	0.020	0.935
19.0	ospananu	0.906	4	21	0.026	0.954
20.5	nurkharl_poncakar	0.902	1	51	0.007	0.889
20.5	levitvas_nazarboh	0.902	1	51	0.007	0.889
22.0	shubevik	0.900	3	34	0.020	0.926
23.0	urbanst6	0.899	1	53	0.007	0.885
24.0	yazykvla	0.889	3	41	0.020	0.911
25.0	kantsmyk_kozlevla	0.888	5	24	0.033	0.948
26.0	barabda1_uherjir3	0.888	2	51	0.013	0.889
27.0	kaslonat_msalhoma	0.886	2	52	0.013	0.887
28.0	horakj31_milyunik	0.876	7	14	0.046	0.970
29.0	divisvi1_douchmar	0.875	8	6	0.052	0.987
30.0	litvidmi_mamaeand	0.868	1	72	0.007	0.844
...



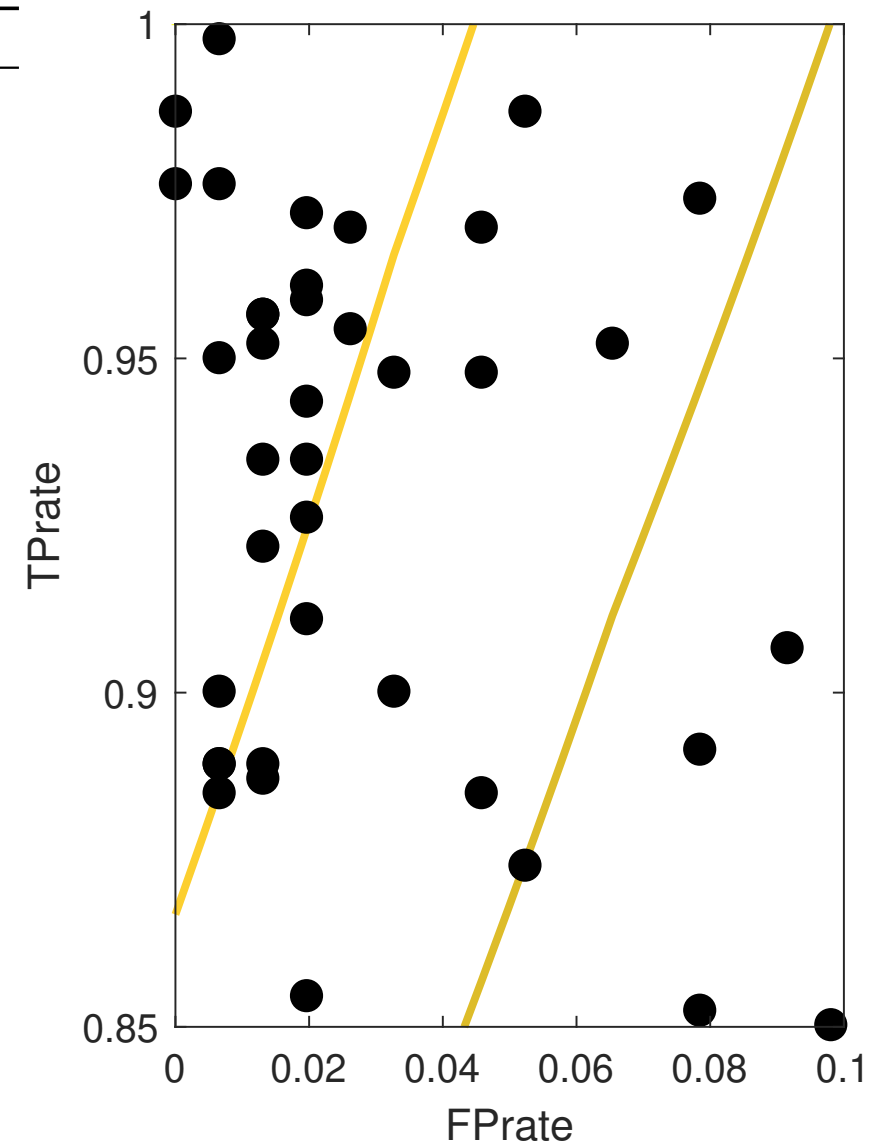
Data 1: výsledky

#	Username	q	FP	FN	FP rate	TP rate
3.0	kyzrondr_rihamir1	0.982	1	1	0.007	0.998
4.0	knoblvit_louckada	0.982	0	11	0.000	0.976
5.0	kovalda2_zribktad	0.966	1	11	0.007	0.976
6.0	pajtaand_sobotma6	0.947	1	23	0.007	0.950
7.5	cezneluk_libalrud	0.937	2	20	0.013	0.957
7.5	blahoja9_bodnafab	0.937	2	20	0.013	0.957
9.0	nejedmat_vonasmic	0.934	2	22	0.013	0.952
10.0	banovvik_belouste	0.933	3	13	0.020	0.972
11.0	mareska9_novacvil	0.925	3	18	0.020	0.961
12.0	saranego	0.924	3	19	0.020	0.959
13.0	weisjose	0.921	2	30	0.013	0.935
14.0	lososjos_stanope3	0.917	4	14	0.026	0.970
15.0	valesja4_zderama1	0.913	3	26	0.020	0.944
16.0	vavermic_ziltjura	0.911	2	36	0.013	0.922
17.0	pelikma3	0.910	1	46	0.007	0.900
18.0	korlavik_korladia	0.906	3	30	0.020	0.935
19.0	ospananu	0.906	4	21	0.026	0.954
20.5	nurkharl_poncakar	0.902	1	51	0.007	0.889
20.5	levitvas_nazarboh	0.902	1	51	0.007	0.889
22.0	shubevik	0.900	3	34	0.020	0.926
23.0	urbanst6	0.899	1	53	0.007	0.885
24.0	yazykvla	0.889	3	41	0.020	0.911
25.0	kantsmyk_kozlevla	0.888	5	24	0.033	0.948
26.0	barabda1_uherjir3	0.888	2	51	0.013	0.889
27.0	kaslonat_msalhoma	0.886	2	52	0.013	0.887
28.0	horakj31_milyunik	0.876	7	14	0.046	0.970
29.0	divisvi1_douchmar	0.875	8	6	0.052	0.987
30.0	litvidmi_mamaeand	0.868	1	72	0.007	0.844
...



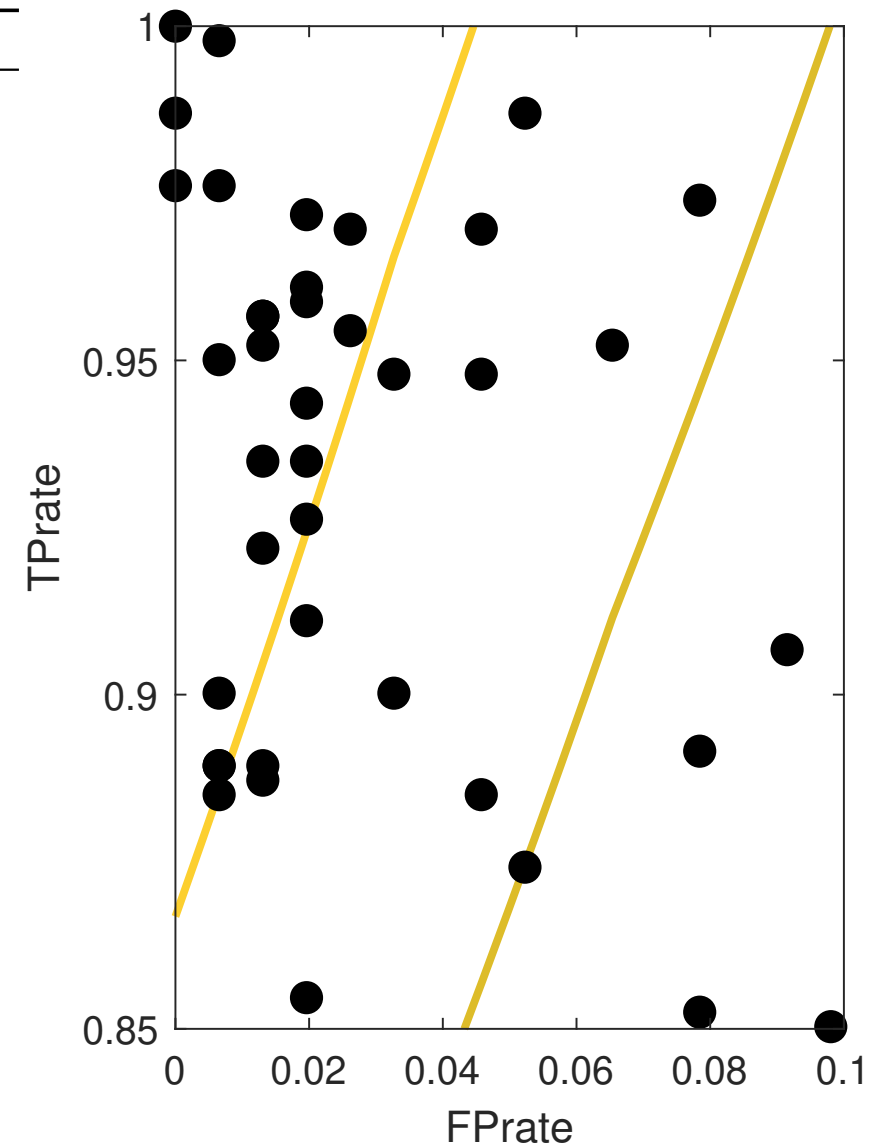
Data 1: výsledky

#	Username	q	FP	FN	FP rate	TP rate
2.0	ondrumat	0.990	0	6	0.000	0.987
3.0	kyzrondr_rihamir1	0.982	1	1	0.007	0.998
4.0	knoblvit_louckada	0.982	0	11	0.000	0.976
5.0	kovalda2_zribktad	0.966	1	11	0.007	0.976
6.0	pajtaand_sobotma6	0.947	1	23	0.007	0.950
7.5	cezneluk_libalrud	0.937	2	20	0.013	0.957
7.5	blahoja9_bodnafab	0.937	2	20	0.013	0.957
9.0	nejedmat_vonasmic	0.934	2	22	0.013	0.952
10.0	banovvik_belouste	0.933	3	13	0.020	0.972
11.0	mareska9_novacvil	0.925	3	18	0.020	0.961
12.0	saranego	0.924	3	19	0.020	0.959
13.0	weisjose	0.921	2	30	0.013	0.935
14.0	lososjos_stanope3	0.917	4	14	0.026	0.970
15.0	valesja4_zderama1	0.913	3	26	0.020	0.944
16.0	vavermic_ziltjura	0.911	2	36	0.013	0.922
17.0	pelikma3	0.910	1	46	0.007	0.900
18.0	korlavik_korladia	0.906	3	30	0.020	0.935
19.0	ospananu	0.906	4	21	0.026	0.954
20.5	nurkharl_poncakar	0.902	1	51	0.007	0.889
20.5	levitvas_nazarboh	0.902	1	51	0.007	0.889
22.0	shubevik	0.900	3	34	0.020	0.926
23.0	urbanst6	0.899	1	53	0.007	0.885
24.0	yazykvla	0.889	3	41	0.020	0.911
25.0	kantsmyk_kozlevla	0.888	5	24	0.033	0.948
26.0	barabda1_uherjir3	0.888	2	51	0.013	0.889
27.0	kaslonat_msalhoma	0.886	2	52	0.013	0.887
28.0	horakj31_milyunik	0.876	7	14	0.046	0.970
29.0	divisvi1_douchmar	0.875	8	6	0.052	0.987
30.0	litvidmi_mamaeand	0.868	1	72	0.007	0.844
...



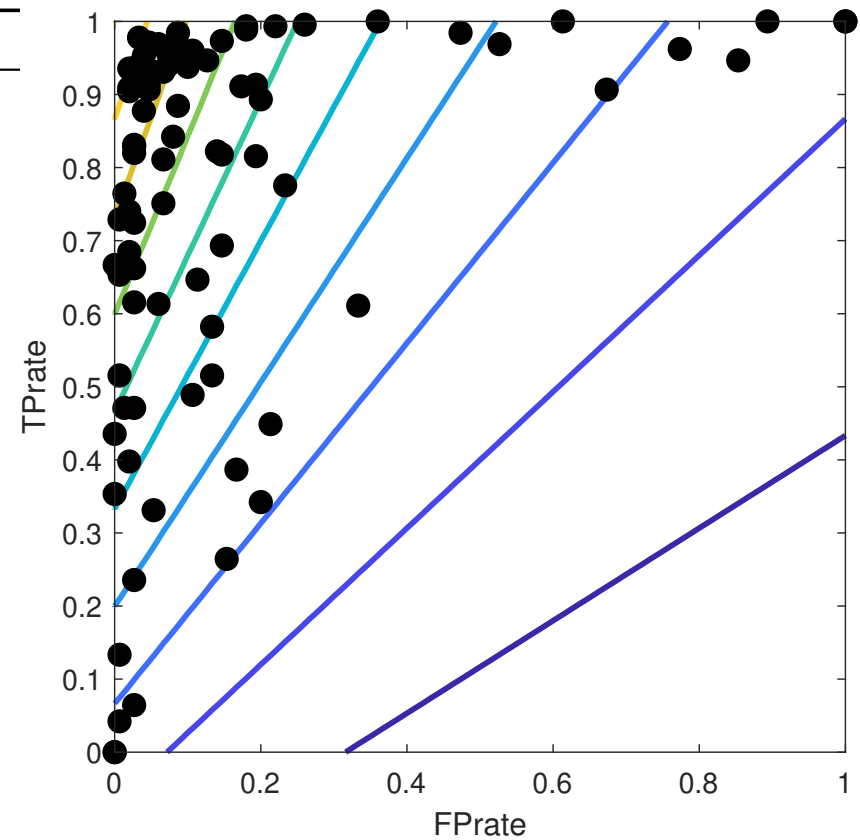
Data 1: výsledky

#	Username	q	FP	FN	FP rate	TP rate
1.0	fraloily_semindan	1.000	0	0	0.000	1.000
2.0	ondrumat	0.990	0	6	0.000	0.987
3.0	kyzrondr_rihamir1	0.982	1	1	0.007	0.998
4.0	knoblvit_louckada	0.982	0	11	0.000	0.976
5.0	kovalda2_zribktad	0.966	1	11	0.007	0.976
6.0	pajtaand_sobotma6	0.947	1	23	0.007	0.950
7.5	cezneluk_libalrud	0.937	2	20	0.013	0.957
7.5	blahoja9_bodnafab	0.937	2	20	0.013	0.957
9.0	nejedmat_vonasmic	0.934	2	22	0.013	0.952
10.0	banovvik_belouste	0.933	3	13	0.020	0.972
11.0	mareska9_novacvil	0.925	3	18	0.020	0.961
12.0	saranego	0.924	3	19	0.020	0.959
13.0	weisjose	0.921	2	30	0.013	0.935
14.0	lososjos_stanope3	0.917	4	14	0.026	0.970
15.0	valesja4_zderama1	0.913	3	26	0.020	0.944
16.0	vavermic_ziltjura	0.911	2	36	0.013	0.922
17.0	pelikma3	0.910	1	46	0.007	0.900
18.0	korlavik_korladia	0.906	3	30	0.020	0.935
19.0	ospananu	0.906	4	21	0.026	0.954
20.5	nurkharl_poncakar	0.902	1	51	0.007	0.889
20.5	levitvas_nazarboh	0.902	1	51	0.007	0.889
22.0	shubevik	0.900	3	34	0.020	0.926
23.0	urbanst6	0.899	1	53	0.007	0.885
24.0	yazykvla	0.889	3	41	0.020	0.911
25.0	kantsmyk_kozlevla	0.888	5	24	0.033	0.948
26.0	barabda1_uherjir3	0.888	2	51	0.013	0.889
27.0	kaslonat_msalhoma	0.886	2	52	0.013	0.887
28.0	horakj31_milyunik	0.876	7	14	0.046	0.970
29.0	divisvil_douchmar	0.875	8	6	0.052	0.987
30.0	litvidmi_mamaeand	0.868	1	72	0.007	0.844
...



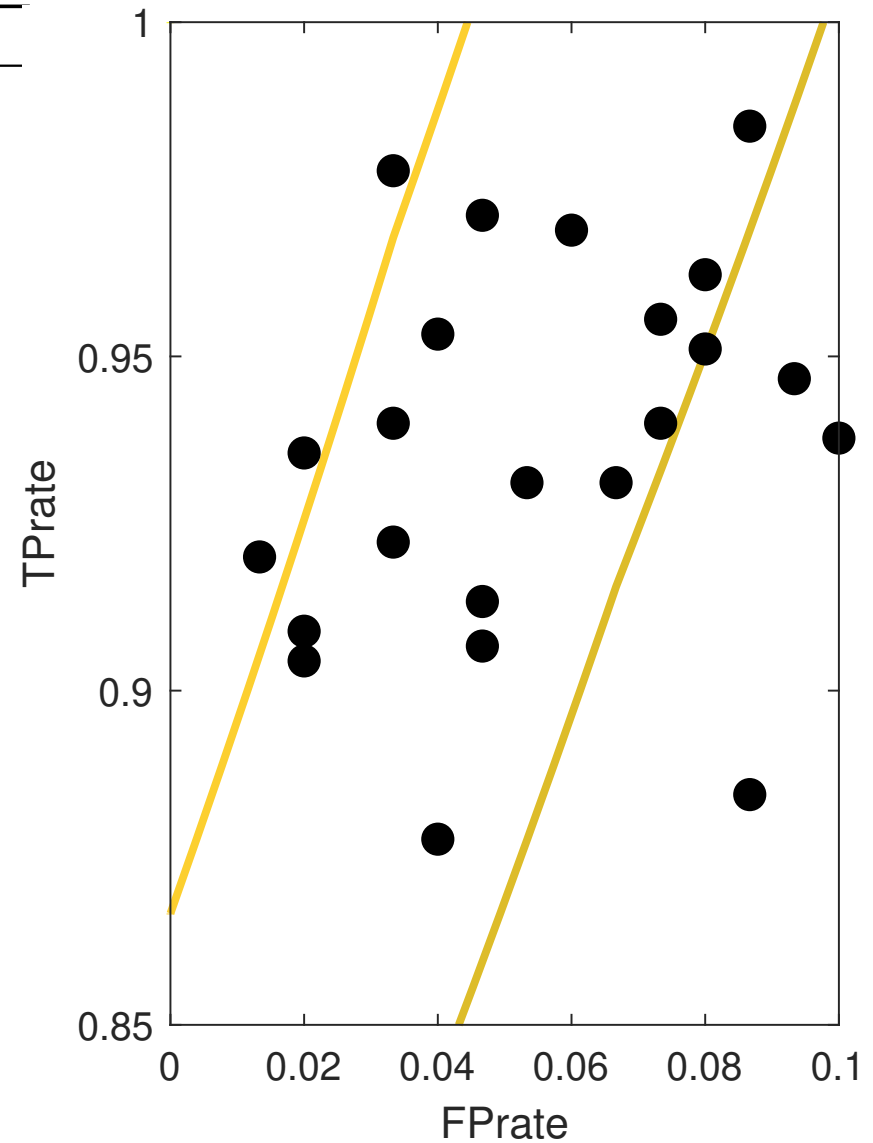
Data 2: výsledky

#	Username	q	FP	FN	FP rate	TP rate
6.0	fraloily_semindan	0.907	5	10	0.033	0.978
7.0	nejedmat_vonasmic	0.906	3	29	0.020	0.936
8.0	yazykvla	0.887	3	41	0.020	0.909
9.0	mamaegeo	0.884	3	43	0.020	0.904
10.0	banovvik_belouste	0.881	5	27	0.033	0.940
11.0	kovalda2_zribktad	0.876	6	21	0.040	0.953
12.0	pajtaand_sobotma6	0.875	7	13	0.047	0.971
13.0	gorevden_lebeddan	0.868	5	35	0.033	0.922
14.0	blahoja9_bodnafab	0.847	9	14	0.060	0.969
15.0	slanizuz_stranste	0.836	7	39	0.047	0.913
16.0	mareska9_novacvi1	0.835	8	31	0.053	0.931
17.0	lososjos_stanope3	0.831	7	42	0.047	0.907
18.0	krausto8_vetrope2	0.824	6	55	0.040	0.878
19.0	valesja4_zderama1	0.818	4	76	0.027	0.831
20.0	golovleo	0.816	4	77	0.027	0.829
21.0	matejm42_mlynatom	0.814	11	20	0.073	0.956
22.0	litvidmi_mamaeand	0.810	10	31	0.067	0.931
23.0	rapandav_ugrynste	0.810	4	81	0.027	0.820
24.0	vavermic_ziltjura	0.809	13	7	0.087	0.984
25.0	liskaji8_liskoreb	0.806	12	17	0.080	0.962
26.0	khursdmy_veselsa1	0.804	11	27	0.073	0.940
27.0	kantsmyk_kozlevla	0.799	12	22	0.080	0.951
28.0	shubevik	0.796	2	106	0.013	0.764
29.0	barabda1_uherjir3	0.783	1	122	0.007	0.729
30.0	saranego	0.774	14	24	0.093	0.947
...



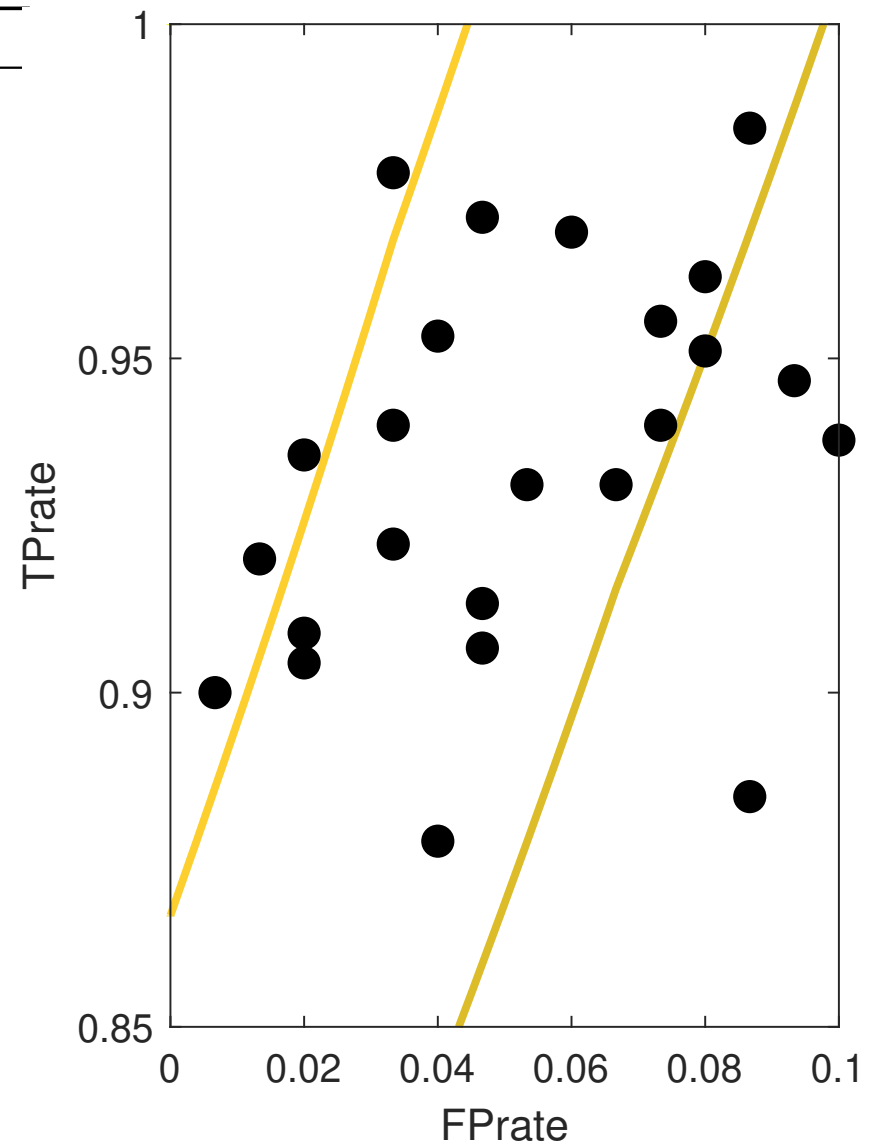
Data 2: výsledky

#	Username	q	FP	FN	FP rate	TP rate
5.0	urbanst6	0.909	2	36	0.013	0.920
6.0	fraloily_semindan	0.907	5	10	0.033	0.978
7.0	nejedmat_vonasmic	0.906	3	29	0.020	0.936
8.0	yazykvla	0.887	3	41	0.020	0.909
9.0	mamaegeo	0.884	3	43	0.020	0.904
10.0	banovvik_belouste	0.881	5	27	0.033	0.940
11.0	kovalda2_zribktad	0.876	6	21	0.040	0.953
12.0	pajtaand_sobotma6	0.875	7	13	0.047	0.971
13.0	gorevden_lebeddan	0.868	5	35	0.033	0.922
14.0	blahoja9_bodnafab	0.847	9	14	0.060	0.969
15.0	slanizuz_stranste	0.836	7	39	0.047	0.913
16.0	mareska9_novacvi1	0.835	8	31	0.053	0.931
17.0	lososjos_stanope3	0.831	7	42	0.047	0.907
18.0	krausto8_vetrope2	0.824	6	55	0.040	0.878
19.0	valesja4_zderama1	0.818	4	76	0.027	0.831
20.0	golovleo	0.816	4	77	0.027	0.829
21.0	matejm42_mlynatom	0.814	11	20	0.073	0.956
22.0	litvidmi_mamaeand	0.810	10	31	0.067	0.931
23.0	rapandav_ugrynste	0.810	4	81	0.027	0.820
24.0	vavermic_ziltjura	0.809	13	7	0.087	0.984
25.0	liskaji8_liskoreb	0.806	12	17	0.080	0.962
26.0	khursdmy_veselsa1	0.804	11	27	0.073	0.940
27.0	kantsmyk_kozlevla	0.799	12	22	0.080	0.951
28.0	shubevik	0.796	2	106	0.013	0.764
29.0	barabda1_uherjir3	0.783	1	122	0.007	0.729
30.0	saranego	0.774	14	24	0.093	0.947
...



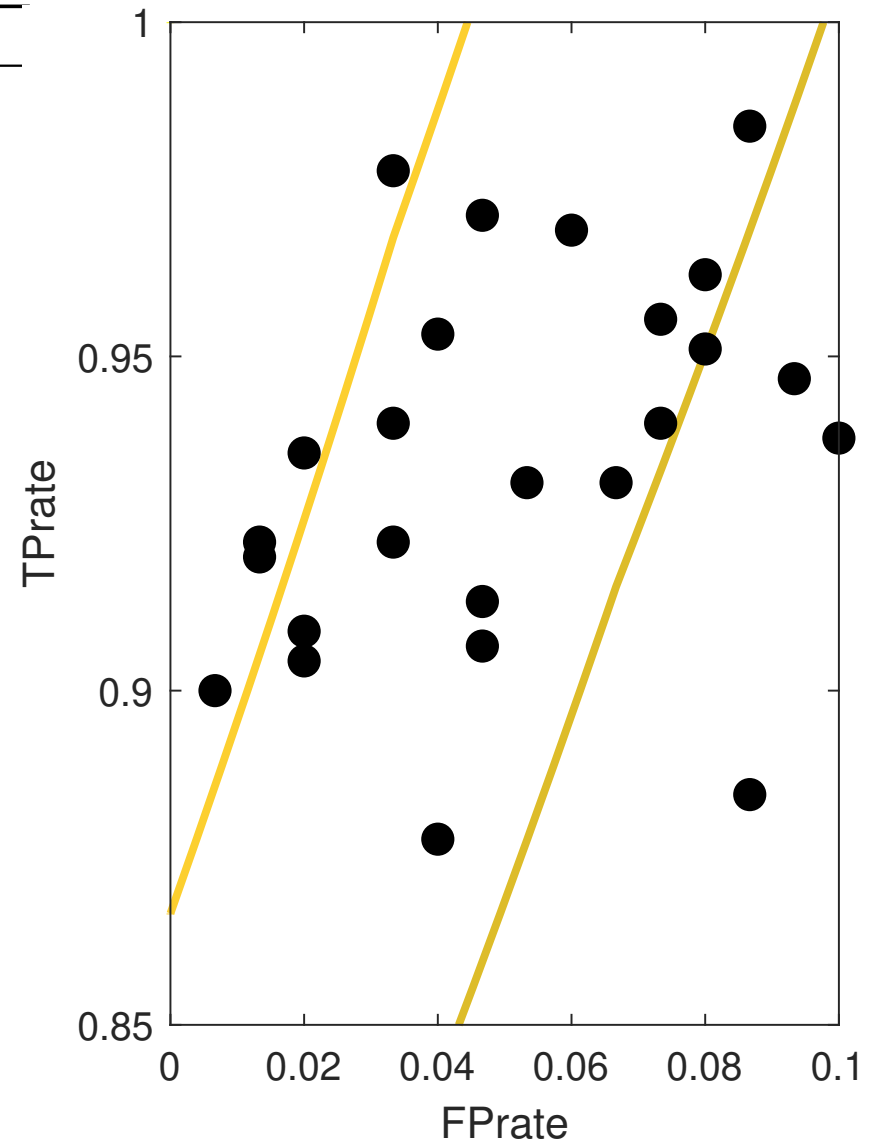
Data 2: výsledky

#	Username	q	FP	FN	FP rate	TP rate
4.0	levitvas_nazarboh	0.910	1	45	0.007	0.900
5.0	urbanst6	0.909	2	36	0.013	0.920
6.0	fraloily_semindan	0.907	5	10	0.033	0.978
7.0	nejedmat_vonasmic	0.906	3	29	0.020	0.936
8.0	yazykvla	0.887	3	41	0.020	0.909
9.0	mamaegeo	0.884	3	43	0.020	0.904
10.0	banovvik_belouste	0.881	5	27	0.033	0.940
11.0	kovalda2_zribktad	0.876	6	21	0.040	0.953
12.0	pajtaand_sobotma6	0.875	7	13	0.047	0.971
13.0	gorevden_lebeddan	0.868	5	35	0.033	0.922
14.0	blahoja9_bodnafab	0.847	9	14	0.060	0.969
15.0	slanizuz_stranste	0.836	7	39	0.047	0.913
16.0	mareska9_novacvi1	0.835	8	31	0.053	0.931
17.0	lososjos_stanope3	0.831	7	42	0.047	0.907
18.0	krausto8_vetrope2	0.824	6	55	0.040	0.878
19.0	valesja4_zderama1	0.818	4	76	0.027	0.831
20.0	golovleo	0.816	4	77	0.027	0.829
21.0	matejm42_mlynatom	0.814	11	20	0.073	0.956
22.0	litvidmi_mamaeand	0.810	10	31	0.067	0.931
23.0	rapandav_ugrynste	0.810	4	81	0.027	0.820
24.0	vavermic_ziltjura	0.809	13	7	0.087	0.984
25.0	liskaji8_liskoreb	0.806	12	17	0.080	0.962
26.0	khursdmy_veselsa1	0.804	11	27	0.073	0.940
27.0	kantsmyk_kozlevla	0.799	12	22	0.080	0.951
28.0	shubevik	0.796	2	106	0.013	0.764
29.0	barabda1_uherjir3	0.783	1	122	0.007	0.729
30.0	saranego	0.774	14	24	0.093	0.947
...



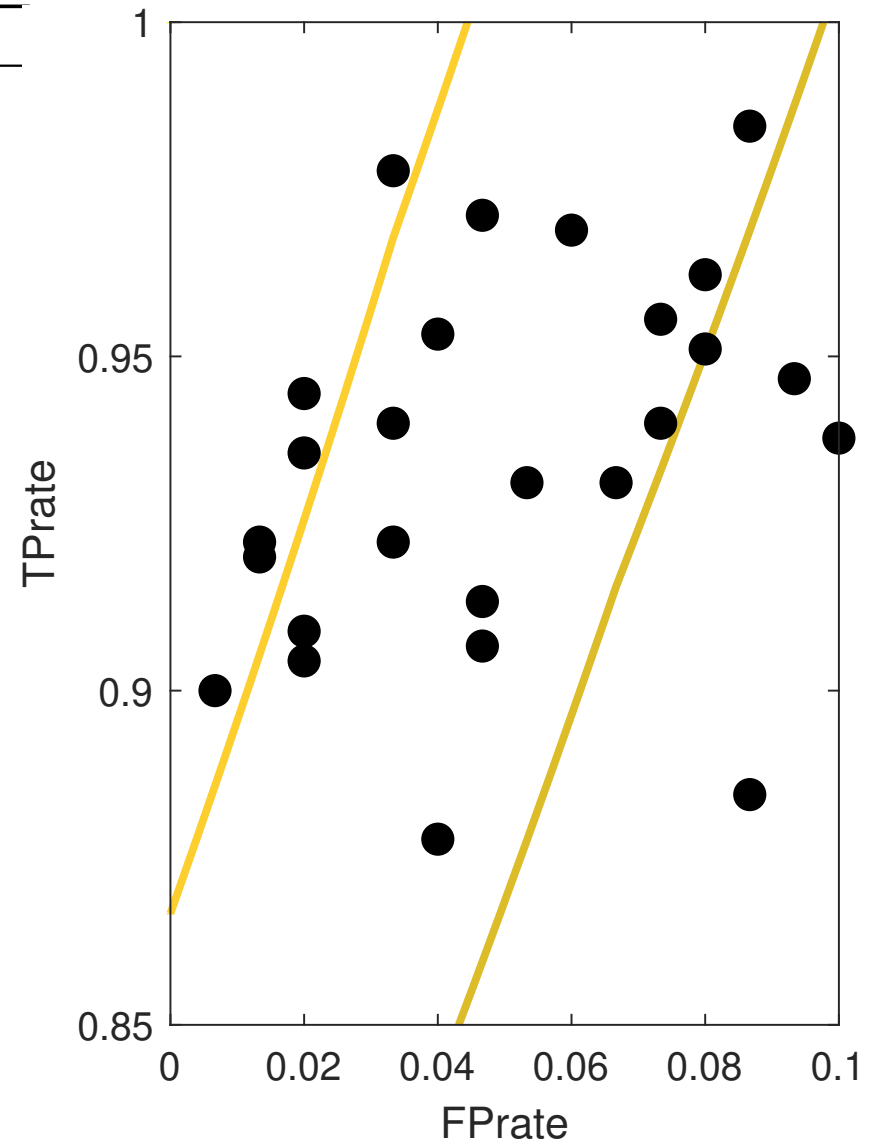
Data 2: výsledky

#	Username	q	FP	FN	FP rate	TP rate
3.0	pelikma3	0.911	2	35	0.013	0.922
4.0	levitvas_nazarboh	0.910	1	45	0.007	0.900
5.0	urbanst6	0.909	2	36	0.013	0.920
6.0	fraloily_semindan	0.907	5	10	0.033	0.978
7.0	nejedmat_vonasmic	0.906	3	29	0.020	0.936
8.0	yazykvla	0.887	3	41	0.020	0.909
9.0	mamaegeo	0.884	3	43	0.020	0.904
10.0	banovvik_belouste	0.881	5	27	0.033	0.940
11.0	kovalda2_zribktad	0.876	6	21	0.040	0.953
12.0	pajtaand_sobotma6	0.875	7	13	0.047	0.971
13.0	gorevden_lebeddan	0.868	5	35	0.033	0.922
14.0	blahoja9_bodnafab	0.847	9	14	0.060	0.969
15.0	slanizuz_stranste	0.836	7	39	0.047	0.913
16.0	mareska9_novacvil	0.835	8	31	0.053	0.931
17.0	lososjos_stanope3	0.831	7	42	0.047	0.907
18.0	krausto8_vetrope2	0.824	6	55	0.040	0.878
19.0	valesja4_zderama1	0.818	4	76	0.027	0.831
20.0	golovleo	0.816	4	77	0.027	0.829
21.0	matejm42_mlynatom	0.814	11	20	0.073	0.956
22.0	litvidmi_mamaeand	0.810	10	31	0.067	0.931
23.0	rapandav_ugrynste	0.810	4	81	0.027	0.820
24.0	vavermic_ziltjura	0.809	13	7	0.087	0.984
25.0	liskaji8_liskoreb	0.806	12	17	0.080	0.962
26.0	khursdmy_veselsa1	0.804	11	27	0.073	0.940
27.0	kantsmyk_kozlevla	0.799	12	22	0.080	0.951
28.0	shubevik	0.796	2	106	0.013	0.764
29.0	barabda1_uherjir3	0.783	1	122	0.007	0.729
30.0	saranego	0.774	14	24	0.093	0.947
...



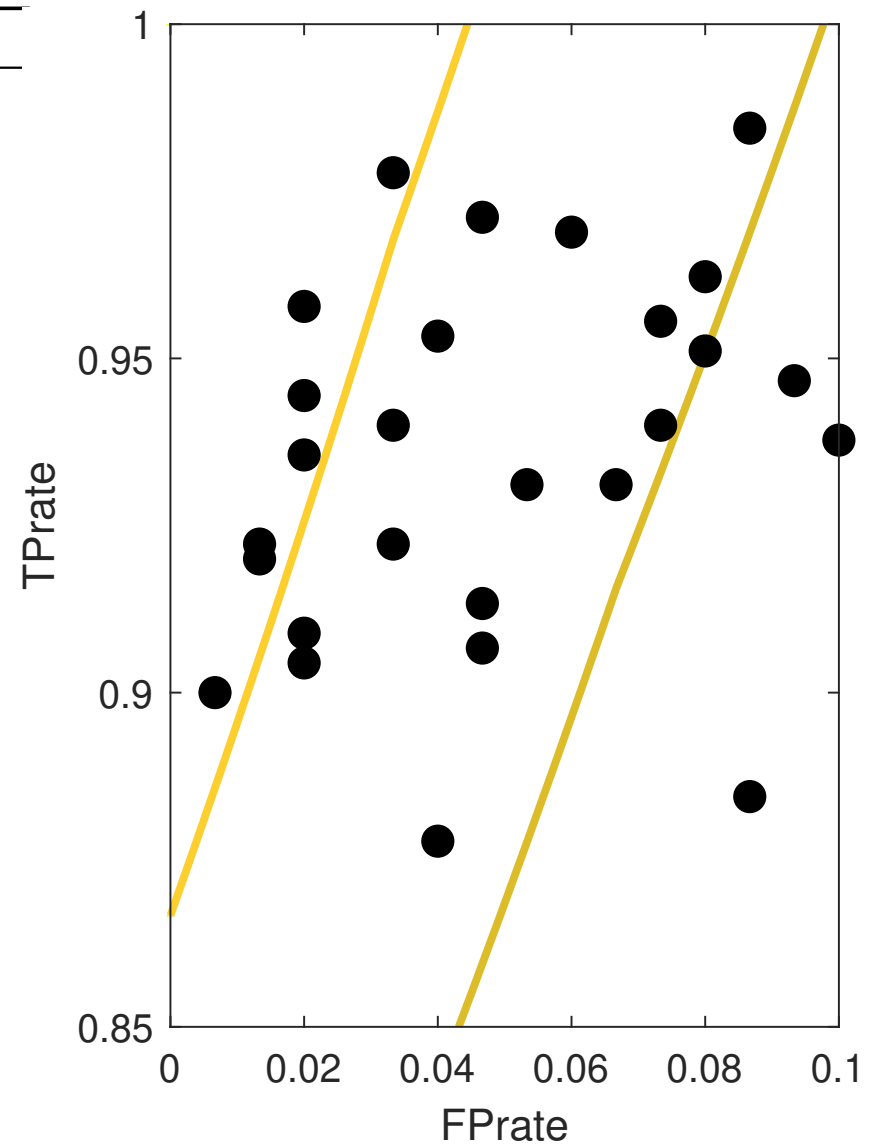
Data 2: výsledky

#	Username	q	FP	FN	FP rate	TP rate
2.0	nurkharl_poncakar	0.912	3	25	0.020	0.944
3.0	pelikma3	0.911	2	35	0.013	0.922
4.0	levitvas_nazarboh	0.910	1	45	0.007	0.900
5.0	urbanst6	0.909	2	36	0.013	0.920
6.0	fraloily_semindan	0.907	5	10	0.033	0.978
7.0	nejedmat_vonasmic	0.906	3	29	0.020	0.936
8.0	yazykvla	0.887	3	41	0.020	0.909
9.0	mamaegeo	0.884	3	43	0.020	0.904
10.0	banovvik_belouste	0.881	5	27	0.033	0.940
11.0	kovalda2_zribktad	0.876	6	21	0.040	0.953
12.0	pajtaand_sobotma6	0.875	7	13	0.047	0.971
13.0	gorevden_lebeddan	0.868	5	35	0.033	0.922
14.0	blahoja9_bodnafab	0.847	9	14	0.060	0.969
15.0	slanizuz_stranste	0.836	7	39	0.047	0.913
16.0	mareska9_novacvil	0.835	8	31	0.053	0.931
17.0	lososjos_stanope3	0.831	7	42	0.047	0.907
18.0	krausto8_vetrope2	0.824	6	55	0.040	0.878
19.0	valesja4_zderama1	0.818	4	76	0.027	0.831
20.0	golovleo	0.816	4	77	0.027	0.829
21.0	matejm42_mlynatom	0.814	11	20	0.073	0.956
22.0	litvidmi_mamaeand	0.810	10	31	0.067	0.931
23.0	rapandav_ugrynste	0.810	4	81	0.027	0.820
24.0	vavermic_ziltjura	0.809	13	7	0.087	0.984
25.0	liskaji8_liskoreb	0.806	12	17	0.080	0.962
26.0	khursdmy_veselsa1	0.804	11	27	0.073	0.940
27.0	kantsmyk_kozlevla	0.799	12	22	0.080	0.951
28.0	shubevik	0.796	2	106	0.013	0.764
29.0	barabda1_uherjir3	0.783	1	122	0.007	0.729
30.0	saranego	0.774	14	24	0.093	0.947
...



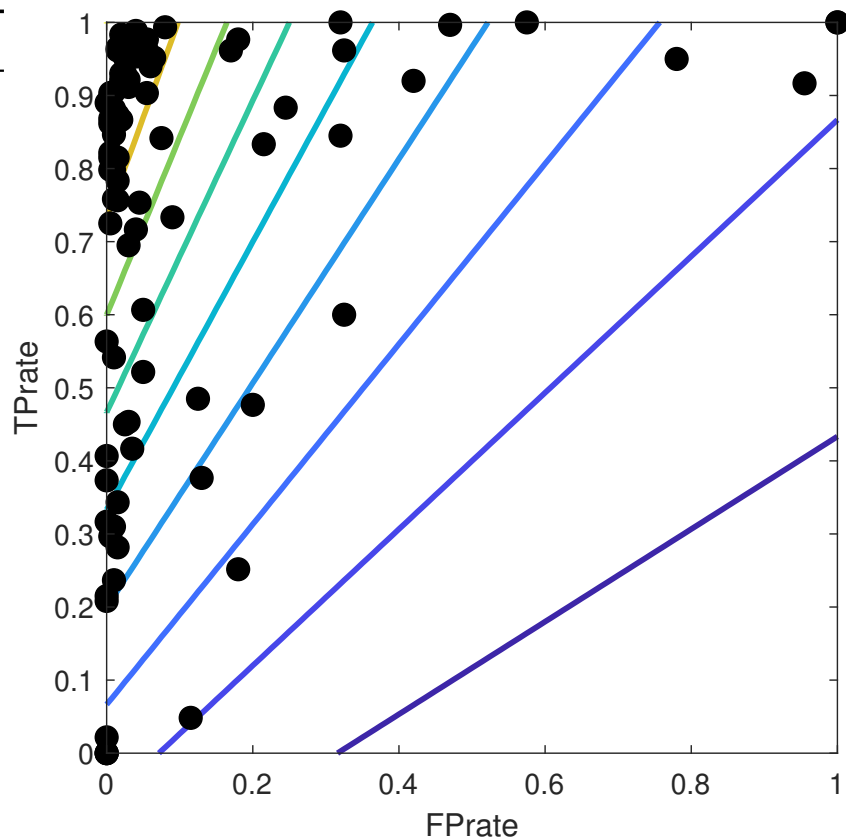
Data 2: výsledky

#	Username	q	FP	FN	FP rate	TP rate
1.0	weisjose	0.922	3	19	0.020	0.958
2.0	nurkharl_poncakar	0.912	3	25	0.020	0.944
3.0	pelikma3	0.911	2	35	0.013	0.922
4.0	levitvas_nazarboh	0.910	1	45	0.007	0.900
5.0	urbanst6	0.909	2	36	0.013	0.920
6.0	fraloily_semindan	0.907	5	10	0.033	0.978
7.0	nejedmat_vonasmic	0.906	3	29	0.020	0.936
8.0	yazykvla	0.887	3	41	0.020	0.909
9.0	mamaegeo	0.884	3	43	0.020	0.904
10.0	banovvik_belouste	0.881	5	27	0.033	0.940
11.0	kovalda2_zribktad	0.876	6	21	0.040	0.953
12.0	pajtaand_sobotma6	0.875	7	13	0.047	0.971
13.0	gorevden_lebeddan	0.868	5	35	0.033	0.922
14.0	blahoja9_bodnafab	0.847	9	14	0.060	0.969
15.0	slanizuz_stranste	0.836	7	39	0.047	0.913
16.0	mareska9_novacvil	0.835	8	31	0.053	0.931
17.0	lososjos_stanope3	0.831	7	42	0.047	0.907
18.0	krausto8_vetrope2	0.824	6	55	0.040	0.878
19.0	valesja4_zderama1	0.818	4	76	0.027	0.831
20.0	golovleo	0.816	4	77	0.027	0.829
21.0	matejm42_mlynatom	0.814	11	20	0.073	0.956
22.0	litvidmi_mamaeand	0.810	10	31	0.067	0.931
23.0	rapandav_ugrynste	0.810	4	81	0.027	0.820
24.0	vavermic_ziltjura	0.809	13	7	0.087	0.984
25.0	liskaji8_liskoreb	0.806	12	17	0.080	0.962
26.0	khursdmy_veselsa1	0.804	11	27	0.073	0.940
27.0	kantsmyk_kozlevla	0.799	12	22	0.080	0.951
28.0	shubevik	0.796	2	106	0.013	0.764
29.0	barabda1_uherjir3	0.783	1	122	0.007	0.729
30.0	saranego	0.774	14	24	0.093	0.947
...



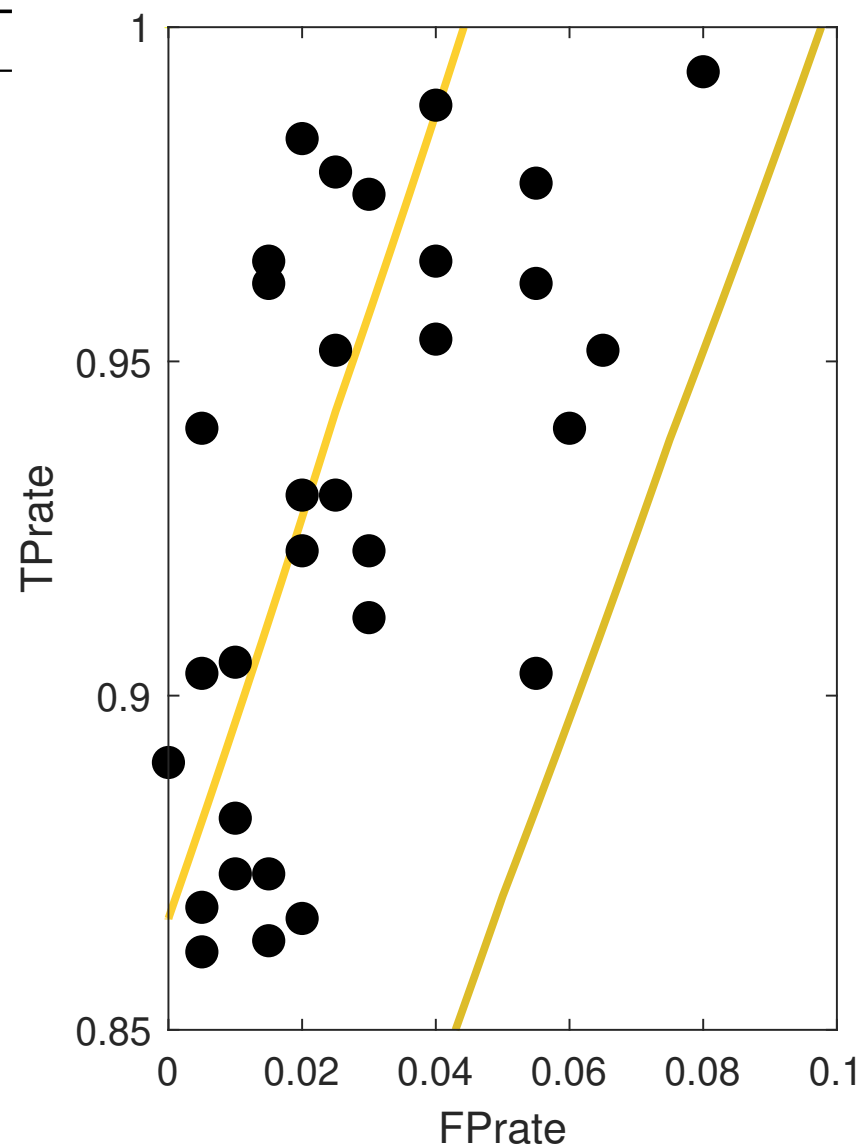
Data 3: výsledky

#	Username	q	FP	FN	FP rate	TP rate
6.0	levitvas_nazarboh	0.940	4	10	0.020	0.983
7.0	vavermic_ziltjura	0.938	3	21	0.015	0.965
8.0	urbanst6	0.936	3	23	0.015	0.962
9.0	pajtaand_sobotma6	0.925	5	13	0.025	0.978
10.0	weisjose	0.917	0	66	0.000	0.890
11.0	knoblvit_louckada	0.916	1	58	0.005	0.903
12.0	yazykvla	0.912	6	15	0.030	0.975
13.0	saranego	0.907	5	29	0.025	0.952
14.0	mareska9_novacvi1	0.906	2	57	0.010	0.905
15.0	golovleo	0.902	4	42	0.020	0.930
16.0	kantsmyk_kozlevla	0.900	8	7	0.040	0.988
17.0	valesja4_zderama1	0.896	4	47	0.020	0.922
18.0	litvidmi_mamaeand	0.891	5	42	0.025	0.930
19.0	kovalda2_zribktad	0.890	1	79	0.005	0.868
20.0	kaslonat_msalhoma	0.889	2	71	0.010	0.882
21.0	lezhnevg	0.885	1	83	0.005	0.862
22.0	cernyad3_lamossta	0.884	8	21	0.040	0.965
23.0	beneson9_capurjos	0.883	2	76	0.010	0.873
24.0	cechdav5_cernyt11	0.876	8	28	0.040	0.953
25.0	erneejan_plevaant	0.875	6	47	0.030	0.922
26.0	ospananu	0.872	3	76	0.015	0.873
27.0	mamaegeo	0.868	6	53	0.030	0.912
28.0	lososjos_stanope3	0.865	3	82	0.015	0.863
29.0	banovvik_belouste	0.863	2	92	0.010	0.847
30.0	pelikma3	0.862	11	14	0.055	0.977
...



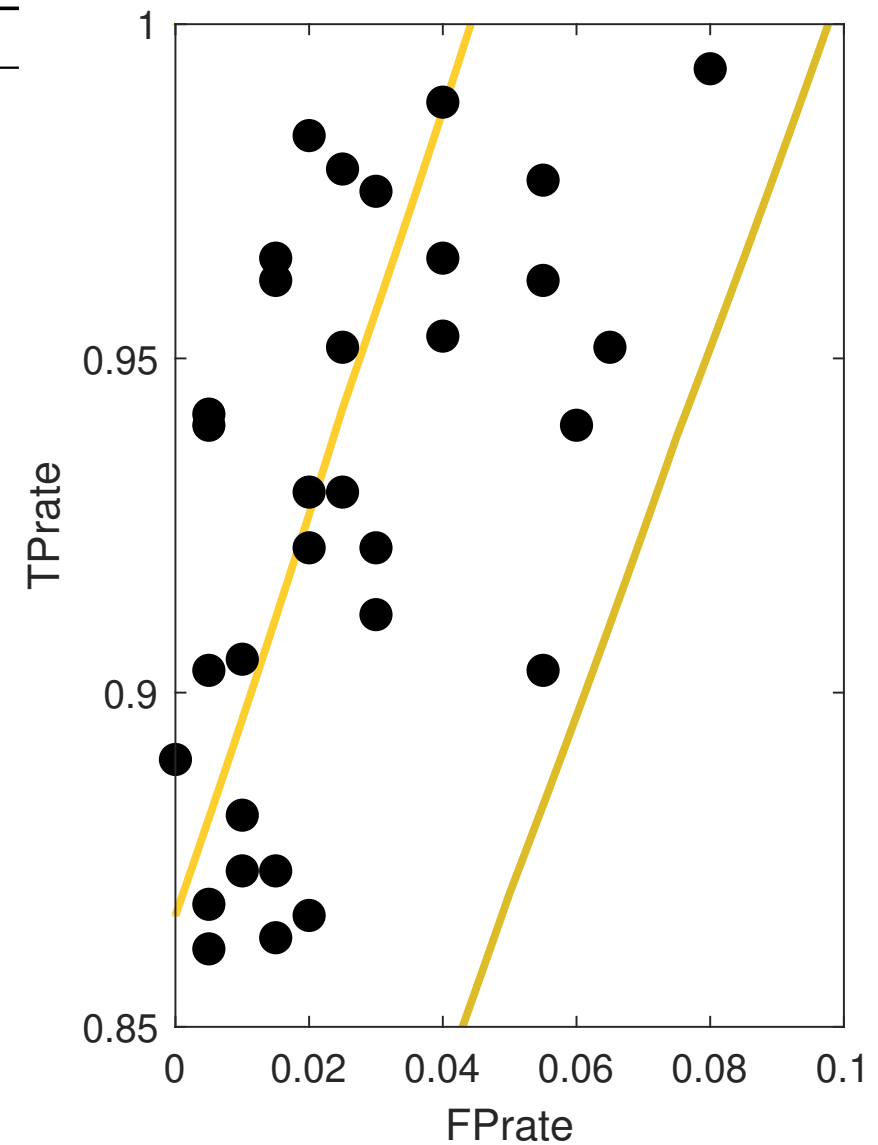
Data 3: výsledky

#	Username	q	FP	FN	FP rate	TP rate
5.0	nejedmat_vonasmic	0.943	1	36	0.005	0.940
6.0	levitvas_nazarboh	0.940	4	10	0.020	0.983
7.0	vavermic_ziltjura	0.938	3	21	0.015	0.965
8.0	urbanst6	0.936	3	23	0.015	0.962
9.0	pajtaand_sobotma6	0.925	5	13	0.025	0.978
10.0	weisjose	0.917	0	66	0.000	0.890
11.0	knoblvit_louckada	0.916	1	58	0.005	0.903
12.0	yazykvla	0.912	6	15	0.030	0.975
13.0	saranego	0.907	5	29	0.025	0.952
14.0	mareska9_novacvi1	0.906	2	57	0.010	0.905
15.0	golovleo	0.902	4	42	0.020	0.930
16.0	kantsmyk_kozlevla	0.900	8	7	0.040	0.988
17.0	valesja4_zderama1	0.896	4	47	0.020	0.922
18.0	litvidmi_mamaeand	0.891	5	42	0.025	0.930
19.0	kovalda2_zribktad	0.890	1	79	0.005	0.868
20.0	kaslonat_msalhoma	0.889	2	71	0.010	0.882
21.0	lezhnevg	0.885	1	83	0.005	0.862
22.0	cernyad3_lamossta	0.884	8	21	0.040	0.965
23.0	beneson9_capurjos	0.883	2	76	0.010	0.873
24.0	cechdav5_cernyt11	0.876	8	28	0.040	0.953
25.0	erneejan_plevaant	0.875	6	47	0.030	0.922
26.0	ospananu	0.872	3	76	0.015	0.873
27.0	mamaegeo	0.868	6	53	0.030	0.912
28.0	lososjos_stanope3	0.865	3	82	0.015	0.863
29.0	banovvik_belouste	0.863	2	92	0.010	0.847
30.0	pelikma3	0.862	11	14	0.055	0.977
...



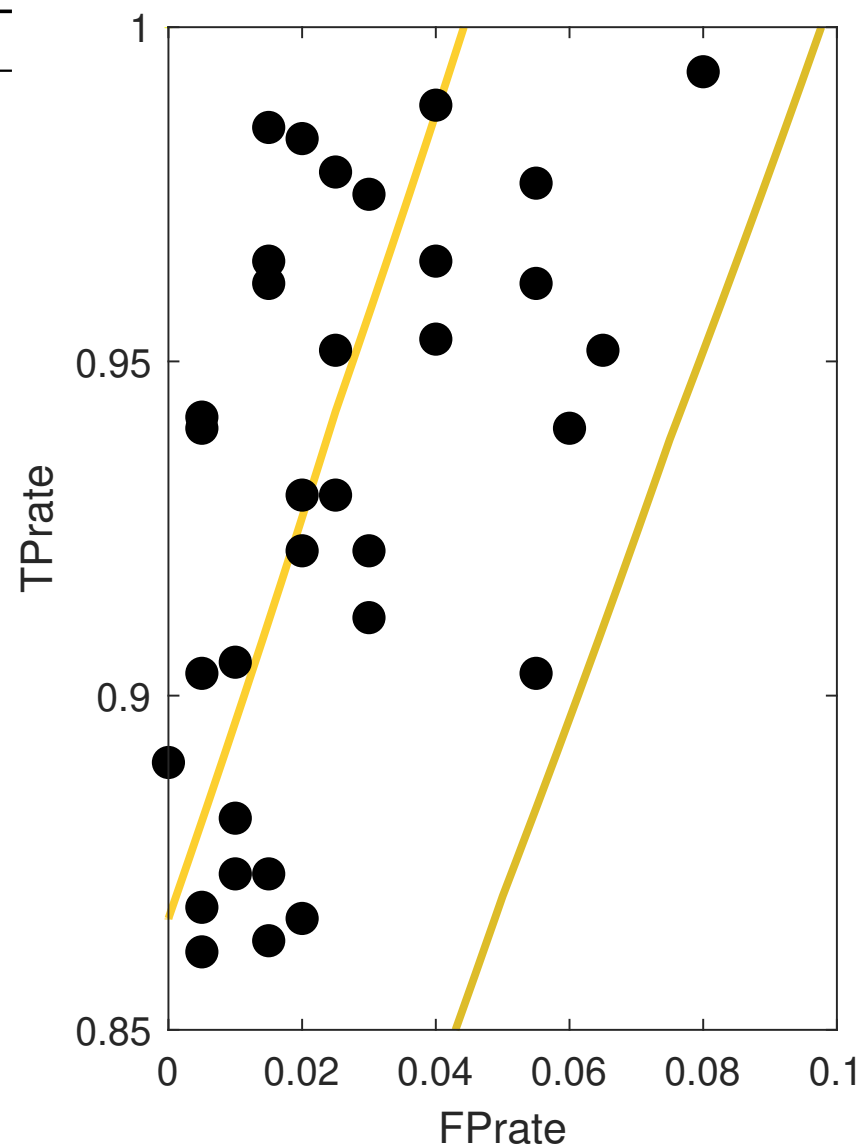
Data 3: výsledky

#	Username	q	FP	FN	FP rate	TP rate
4.0	barabda1_uherjir3	0.944	1	35	0.005	0.942
5.0	nejedmat_vonasmic	0.943	1	36	0.005	0.940
6.0	levitvas_nazarboh	0.940	4	10	0.020	0.983
7.0	vavermic_ziltjura	0.938	3	21	0.015	0.965
8.0	urbanst6	0.936	3	23	0.015	0.962
9.0	pajtaand_sobotma6	0.925	5	13	0.025	0.978
10.0	weisjose	0.917	0	66	0.000	0.890
11.0	knoblvit_louckada	0.916	1	58	0.005	0.903
12.0	yazykvla	0.912	6	15	0.030	0.975
13.0	saranego	0.907	5	29	0.025	0.952
14.0	mareska9_novacvi1	0.906	2	57	0.010	0.905
15.0	golovleo	0.902	4	42	0.020	0.930
16.0	kantsmyk_kozlevla	0.900	8	7	0.040	0.988
17.0	valesja4_zderama1	0.896	4	47	0.020	0.922
18.0	litvidmi_mamaeand	0.891	5	42	0.025	0.930
19.0	kovalda2_zribktad	0.890	1	79	0.005	0.868
20.0	kaslonat_msalhoma	0.889	2	71	0.010	0.882
21.0	lezhnevg	0.885	1	83	0.005	0.862
22.0	cernyad3_lamossta	0.884	8	21	0.040	0.965
23.0	beneson9_capurjos	0.883	2	76	0.010	0.873
24.0	cechdav5_cernyt11	0.876	8	28	0.040	0.953
25.0	erneejan_plevaant	0.875	6	47	0.030	0.922
26.0	ospananu	0.872	3	76	0.015	0.873
27.0	mamaegeo	0.868	6	53	0.030	0.912
28.0	lososjos_stanope3	0.865	3	82	0.015	0.863
29.0	banovvik_belouste	0.863	2	92	0.010	0.847
30.0	pelikma3	0.862	11	14	0.055	0.977
...



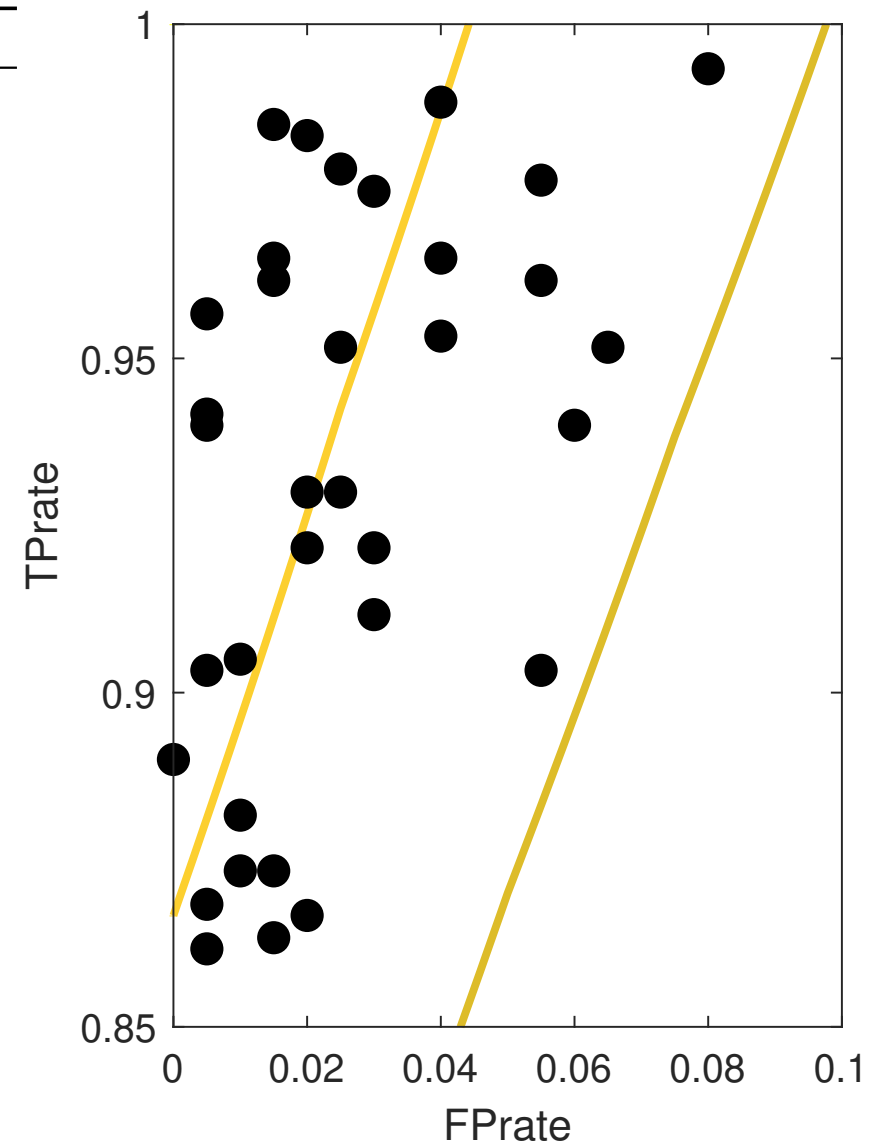
Data 3: výsledky

#	Username	q	FP	FN	FP rate	TP rate
3.0	fraloily_semindan	0.953	3	9	0.015	0.985
4.0	barabda1_uherjir3	0.944	1	35	0.005	0.942
5.0	nejedmat_vonasmic	0.943	1	36	0.005	0.940
6.0	levitvas_nazarboh	0.940	4	10	0.020	0.983
7.0	vavermic_ziltjura	0.938	3	21	0.015	0.965
8.0	urbanst6	0.936	3	23	0.015	0.962
9.0	pajtaand_sobotma6	0.925	5	13	0.025	0.978
10.0	weisjose	0.917	0	66	0.000	0.890
11.0	knoblvit_louckada	0.916	1	58	0.005	0.903
12.0	yazykvla	0.912	6	15	0.030	0.975
13.0	saranego	0.907	5	29	0.025	0.952
14.0	mareska9_novacvil	0.906	2	57	0.010	0.905
15.0	golovleo	0.902	4	42	0.020	0.930
16.0	kantsmyk_kozlevla	0.900	8	7	0.040	0.988
17.0	valesja4_zderama1	0.896	4	47	0.020	0.922
18.0	litvidmi_mamaeand	0.891	5	42	0.025	0.930
19.0	kovalda2_zribktad	0.890	1	79	0.005	0.868
20.0	kaslonat_msalhoma	0.889	2	71	0.010	0.882
21.0	lezhnevg	0.885	1	83	0.005	0.862
22.0	cernyad3_lamossta	0.884	8	21	0.040	0.965
23.0	beneson9_capurjos	0.883	2	76	0.010	0.873
24.0	cechdav5_cernyt11	0.876	8	28	0.040	0.953
25.0	erneejan_plevaant	0.875	6	47	0.030	0.922
26.0	ospananu	0.872	3	76	0.015	0.873
27.0	mamaegeo	0.868	6	53	0.030	0.912
28.0	lososjos_stanope3	0.865	3	82	0.015	0.863
29.0	banovvik_belouste	0.863	2	92	0.010	0.847
30.0	pelikma3	0.862	11	14	0.055	0.977
...



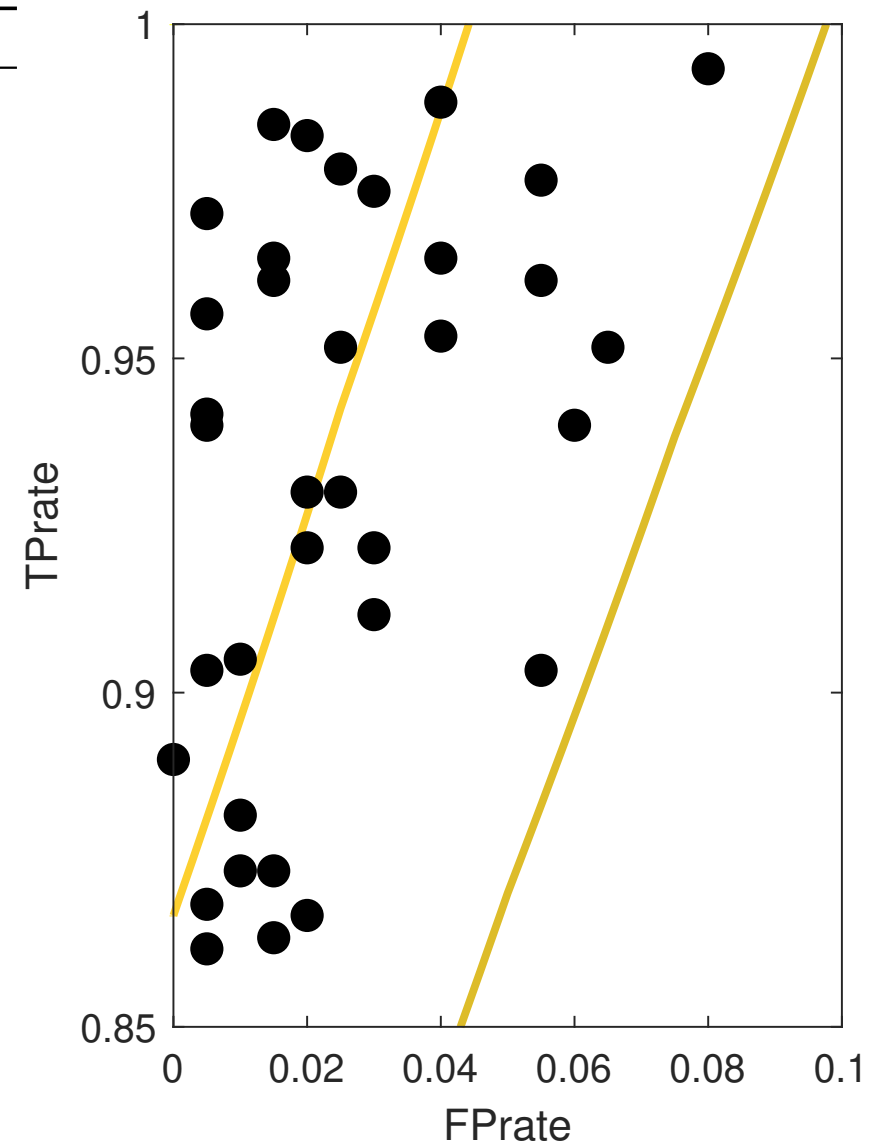
Data 3: výsledky

#	Username	q	FP	FN	FP rate	TP rate
2.0	blahoja9_bodnafab	0.956	1	26	0.005	0.957
3.0	fraloily_semindan	0.953	3	9	0.015	0.985
4.0	barabda1_uherjir3	0.944	1	35	0.005	0.942
5.0	nejedmat_vonasmic	0.943	1	36	0.005	0.940
6.0	levitvas_nazarboh	0.940	4	10	0.020	0.983
7.0	vavermic_ziltjura	0.938	3	21	0.015	0.965
8.0	urbanst6	0.936	3	23	0.015	0.962
9.0	pajtaand_sobotma6	0.925	5	13	0.025	0.978
10.0	weisjose	0.917	0	66	0.000	0.890
11.0	knoblvit_louckada	0.916	1	58	0.005	0.903
12.0	yazykvla	0.912	6	15	0.030	0.975
13.0	saranego	0.907	5	29	0.025	0.952
14.0	mareska9_novacvi1	0.906	2	57	0.010	0.905
15.0	golovleo	0.902	4	42	0.020	0.930
16.0	kantsmyk_kozlevla	0.900	8	7	0.040	0.988
17.0	valesja4_zderama1	0.896	4	47	0.020	0.922
18.0	litvidmi_mamaeand	0.891	5	42	0.025	0.930
19.0	kovalda2_zribktad	0.890	1	79	0.005	0.868
20.0	kaslonat_msalhoma	0.889	2	71	0.010	0.882
21.0	lezhnevg	0.885	1	83	0.005	0.862
22.0	cernyad3_lamossta	0.884	8	21	0.040	0.965
23.0	beneson9_capurjos	0.883	2	76	0.010	0.873
24.0	cechdav5_cernyt11	0.876	8	28	0.040	0.953
25.0	erneejan_plevaant	0.875	6	47	0.030	0.922
26.0	ospananu	0.872	3	76	0.015	0.873
27.0	mamaegeo	0.868	6	53	0.030	0.912
28.0	lososjos_stanope3	0.865	3	82	0.015	0.863
29.0	banovvik_belouste	0.863	2	92	0.010	0.847
30.0	pelikma3	0.862	11	14	0.055	0.977
...



Data 3: výsledky

#	Username	q	FP	FN	FP rate	TP rate
1.0	pivrndav_pospika6	0.967	1	17	0.005	0.972
2.0	blahoja9_bodnafab	0.956	1	26	0.005	0.957
3.0	fraloily_semindan	0.953	3	9	0.015	0.985
4.0	barabda1_uherjir3	0.944	1	35	0.005	0.942
5.0	nejedmat_vonasmic	0.943	1	36	0.005	0.940
6.0	levitvas_nazarboh	0.940	4	10	0.020	0.983
7.0	vavermic_ziltjura	0.938	3	21	0.015	0.965
8.0	urbanst6	0.936	3	23	0.015	0.962
9.0	pajtaand_sobotma6	0.925	5	13	0.025	0.978
10.0	weisjose	0.917	0	66	0.000	0.890
11.0	knoblvit_louckada	0.916	1	58	0.005	0.903
12.0	yazykvla	0.912	6	15	0.030	0.975
13.0	saranego	0.907	5	29	0.025	0.952
14.0	mareska9_novacvi1	0.906	2	57	0.010	0.905
15.0	golovleo	0.902	4	42	0.020	0.930
16.0	kantsmyk_kozlevla	0.900	8	7	0.040	0.988
17.0	valesja4_zderama1	0.896	4	47	0.020	0.922
18.0	litvidmi_mamaeand	0.891	5	42	0.025	0.930
19.0	kovalda2_zribktad	0.890	1	79	0.005	0.868
20.0	kaslonat_msalhoma	0.889	2	71	0.010	0.882
21.0	lezhnevg	0.885	1	83	0.005	0.862
22.0	cernyad3_lamossta	0.884	8	21	0.040	0.965
23.0	beneson9_capurjos	0.883	2	76	0.010	0.873
24.0	cechdav5_cernyt11	0.876	8	28	0.040	0.953
25.0	erneejan_plevaant	0.875	6	47	0.030	0.922
26.0	ospananu	0.872	3	76	0.015	0.873
27.0	mamaegeo	0.868	6	53	0.030	0.912
28.0	lososjos_stanope3	0.865	3	82	0.015	0.863
29.0	banovvik_belouste	0.863	2	92	0.010	0.847
30.0	pelikma3	0.862	11	14	0.055	0.977
...



Porovnání výsledků na 3 sadách dat

Data 1:

#	Username	q
1.0	fral_semi	1.000
2.0	ondrumat	0.990
3.0	kyzr_riha	0.982
4.0	knob_louc	0.982
5.0	kova_zrib	0.966
6.0	pajt_sobo	0.947
7.5	blah_bodn	0.937
7.5	cezn_liba	0.937
9.0	neje_vona	0.934
10.0	bano_belo	0.933
11.0	mare_nova	0.925
12.0	saranego	0.924
13.0	weisjose	0.921
14.0	loso_stan	0.917
15.0	vale_zder	0.913
16.0	vave_zilt	0.911
17.0	pelikma3	0.910
18.0	korl_korl	0.906
19.0	ospananu	0.906
20.5	levi_naza	0.902
20.5	nurk_ponc	0.902
22.0	shubevik	0.900
23.0	urbanst6	0.899
24.0	yazykvla	0.889
25.0	kant_kozl	0.888
...	others	...

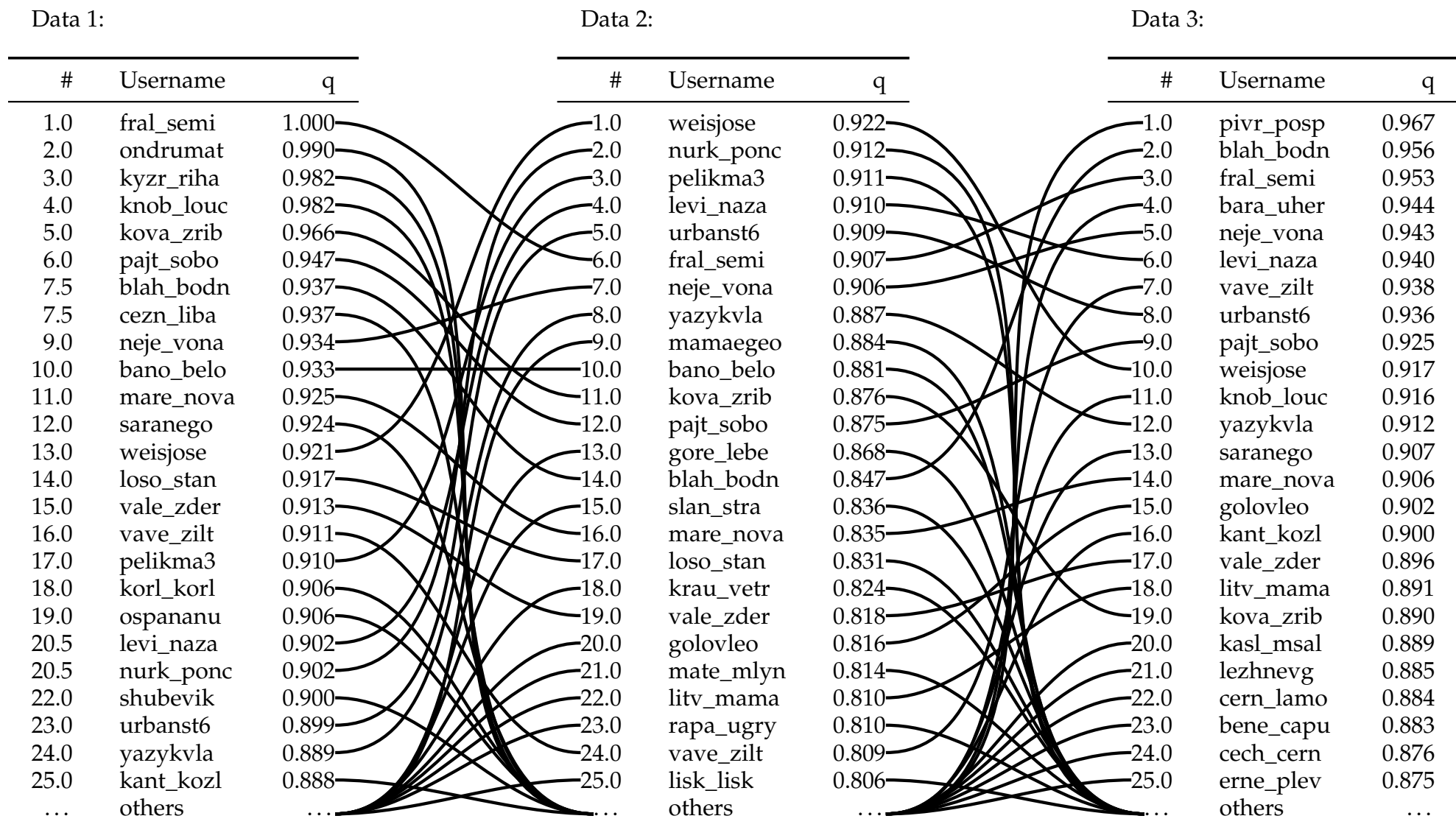
Data 2:

#	Username	q
1.0	weisjose	0.922
2.0	nurk_ponc	0.912
3.0	pelikma3	0.911
4.0	levi_naza	0.910
5.0	urbanst6	0.909
6.0	fral_semi	0.907
7.0	neje_vona	0.906
8.0	yazykvla	0.887
9.0	mamaegeo	0.884
10.0	bano_belo	0.881
11.0	kova_zrib	0.876
12.0	pajt_sobo	0.875
13.0	gore_lebe	0.868
14.0	blah_bodn	0.847
15.0	slan_stra	0.836
16.0	mare_nova	0.835
17.0	loso_stan	0.831
18.0	krau_vetr	0.824
19.0	vale_zder	0.818
20.0	golovleo	0.816
21.0	mate_mlyn	0.814
22.0	litv_mama	0.810
23.0	rapa_ugry	0.810
24.0	vave_zilt	0.809
25.0	lisk_lisk	0.806
...	others	...

Data 3:

#	Username	q
1.0	pivr_posp	0.967
2.0	blah_bodn	0.956
3.0	fral_semi	0.953
4.0	bara_uher	0.944
5.0	neje_vona	0.943
6.0	levi_naza	0.940
7.0	vave_zilt	0.938
8.0	urbanst6	0.936
9.0	pajt_sobo	0.925
10.0	weisjose	0.917
11.0	knob_louc	0.916
12.0	yazykvla	0.912
13.0	saranego	0.907
14.0	mare_nova	0.906
15.0	golovleo	0.902
16.0	kant_kozl	0.900
17.0	vale_zder	0.896
18.0	litv_mama	0.891
19.0	kova_zrib	0.890
20.0	kasl_msar	0.889
21.0	lezhnevg	0.885
22.0	cern_lamo	0.884
23.0	bene_capu	0.883
24.0	cech_cern	0.876
25.0	erne_plev	0.875
...	others	...

Porovnání výsledků na 3 sadách dat





Vítězné filtry

Hodnocení binárních
klasifikátorů
a vizualizace jejich
úspěšnosti

Spam filtr

- Trocha suché statistiky
- 3 datové sady
- Bodové hodnocení
- Data 1: výsledky
- Data 2: výsledky
- Data 3: výsledky
- Porovnání výsledků na 3 sadách dat
- **Vítězné filtry**

Shrnutí

Dotazy?

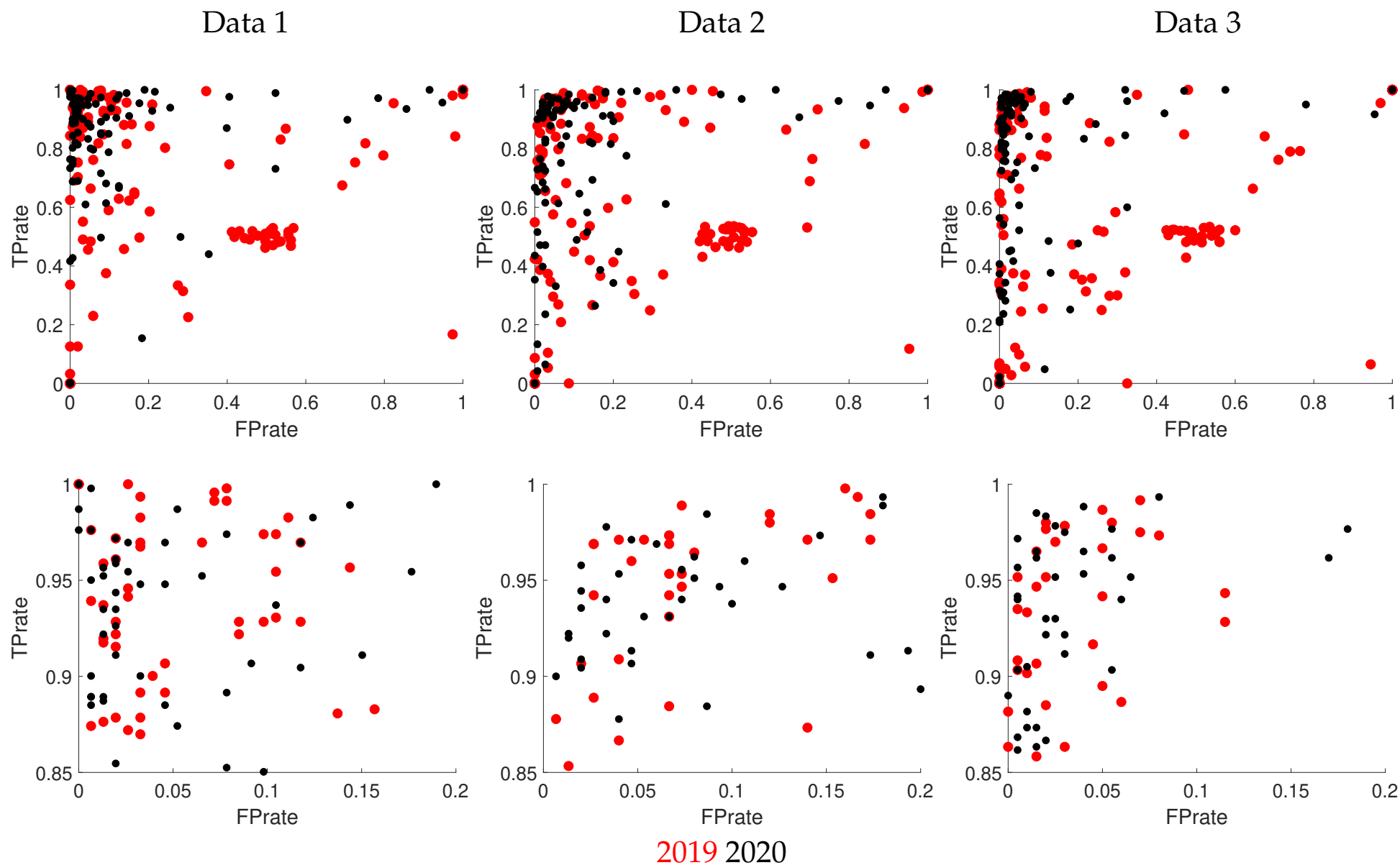
1. Ilya Fralou, Danil Semin (Pořadí: $1 + 6 + 3 = 10$)
2. Matěj Nejedlý, Michal Vonášek (Pořadí: $9 + 7 + 5 = 21$)
3. Jan Blahoš, Fabián Bodnár (Pořadí: $7.5 + 14 + 2 = 23.5$)

Rok	1.	2.	3.
2020	10 (1+6+3)	21 (9+7+5)	23.5 (7.5+14+2)
2019	22 (10+5+7)	33 (2+14+17)	37 (13+21+3)
2018	9 (3+3+3)	12 (3+1+8)	12 (3+4+5)
2017	5 (1+3+1)	8 (1+2+5)	12 (5+1+6)
2016	22 (4+3+15)	31 (22+8+1)	35 (15+8+12)
2015	3 (1+1+1)	6 (1+3+2)	8 (1+4+3)
2014	6 (1+3+2)	12 (1+10+1)	19 (7+1+11)
2013	4 (1+1+2)	6 (1+2+3)	10 (1+4+5)
2012	4 (1+2+1)	11 (7+1+3)	24 (5+14+5)
2011	4 (2+1+1)	9 (3+4+2)	11 (5+3+3)



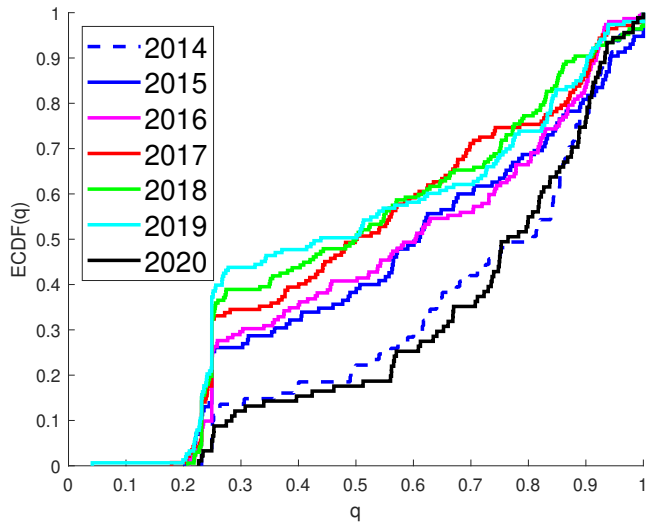
Shrnutí

Porovnání s předchozím rokem

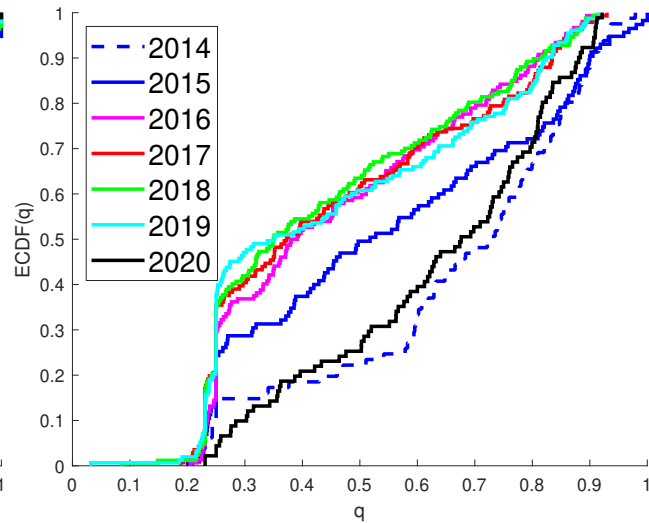


Porovnání s minulými lety (ECDF)

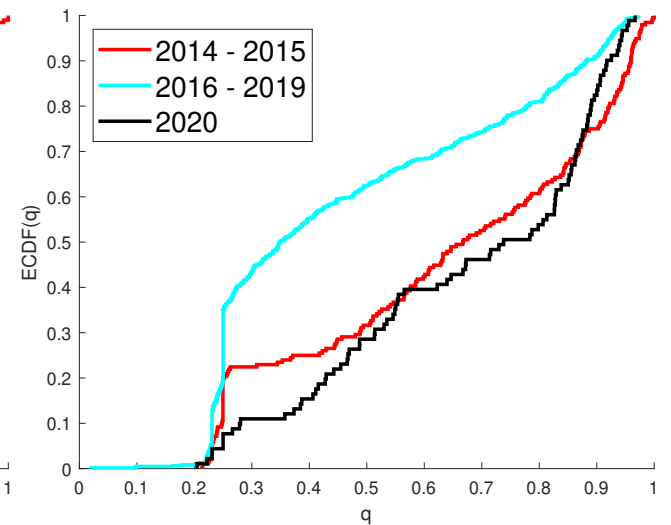
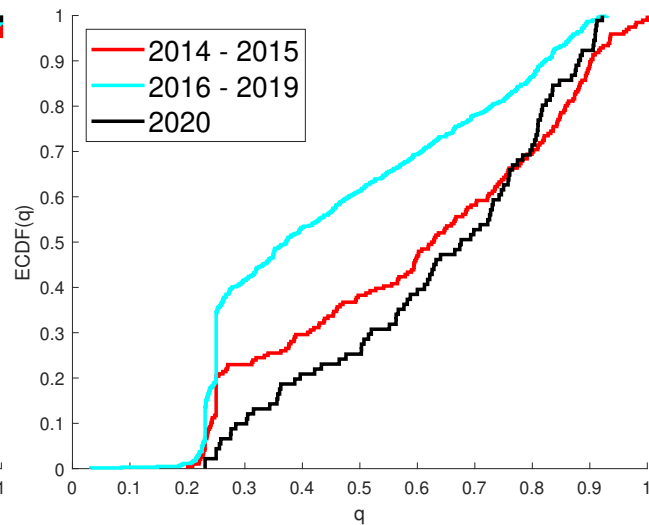
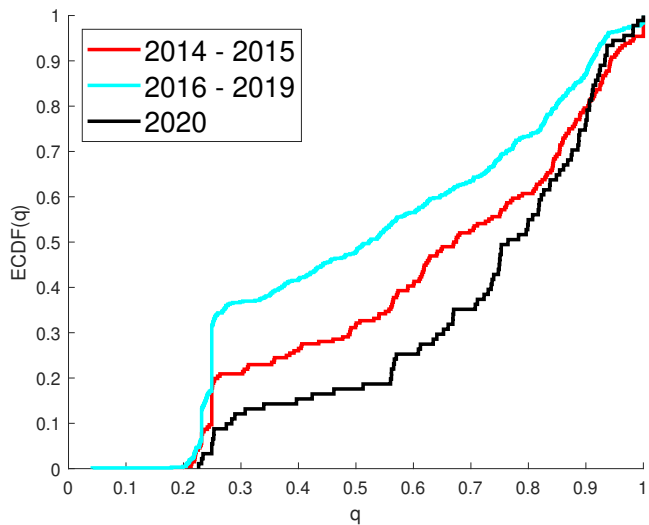
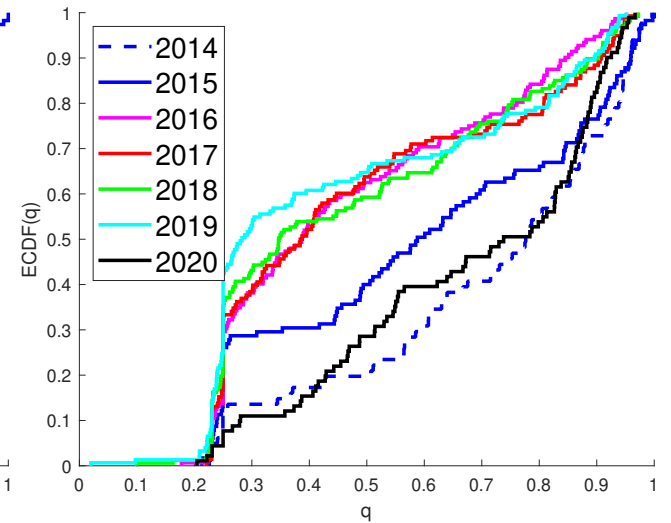
Data 1



Data 2 (rozdílná od 2016)

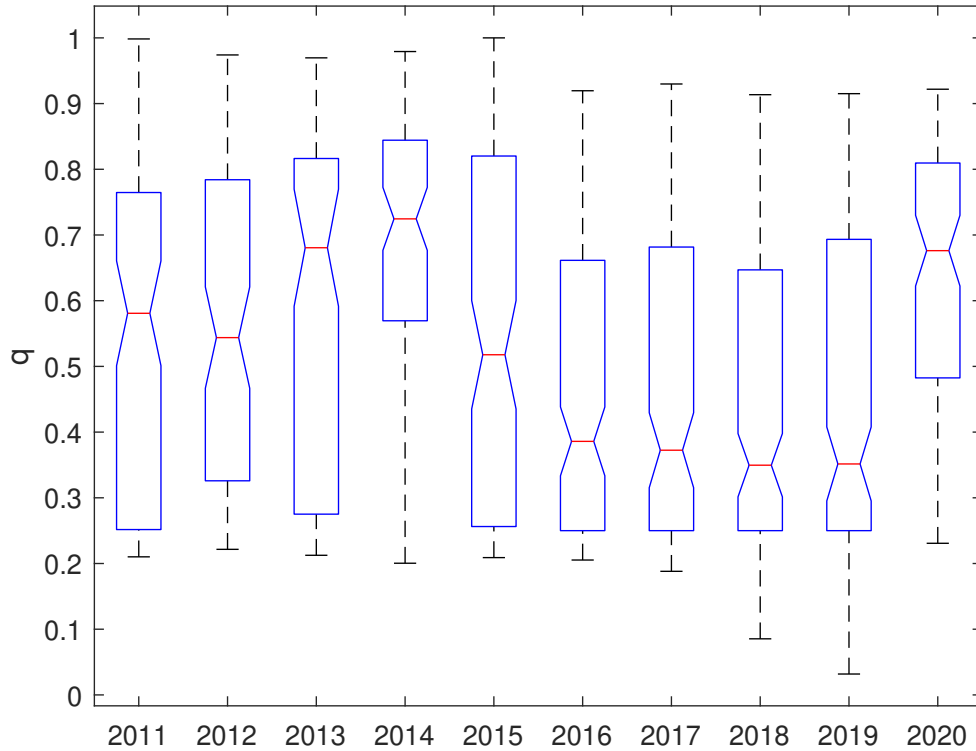


Data 3 (rozdílná od 2016)

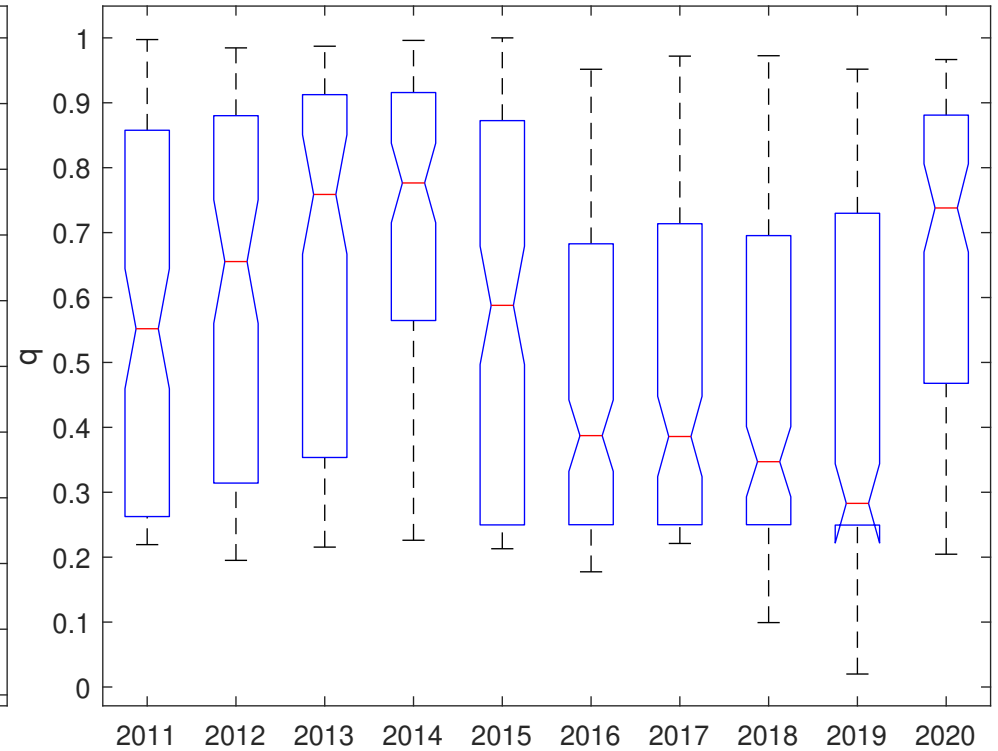


Porovnání s minulými lety (Box)

Contest 2



Contest 3





Pozorování a závěry

Typy filtrů:

Rok	Triviální	Pravidlo	Složitější	Učení
2019	28 %	7 %	5 %	60 %
2018	27 %	8 %	7 %	58 %
2017	25 %	6 %	5 %	64 %

Hodnocení binárních klasifikátorů a vizualizace jejich úspěšnosti

Spam filtr

Shrnutí

- Porovnání s předchozím rokem
- **Pozorování a závěry**
- Na jaké další předměty se můžete těšit?

Dotazy?

Cesty, které k cíli nevedly:

- náhodné rozhodování
- konstantní rozhodování: “klasifikuj všechno jako spam/ham”
- použití statického seznamu zakázaných slov nebo symbolů

Podmínky úspěchu:

- četnosti slov v hamech a spamech jako základ
- výběr diskriminujících slov (která silně podporují/vyvracejí hypotézu SPAM)
- “vyladěný” způsob kombinace důkazů pro a proti hypotéze SPAM
- když si nejsm jistý, klasifikovat spíše jako ham



Na jaké další předměty se můžete těšit?

Hodnocení binárních
klasifikátorů
a vizualizace jejich
úspěšnosti

Spam filtr

Shrnutí

- Porovnání s předchozím rokem
- Pozorování a závěry
- Na jaké další předměty se můžete těšit?

Dotazy?

Bakalářská etapa:

- B4B33ALG: Algoritmizace – Marko Genyk-Berezovskyj
- B0B01PST: Pravděpodobnost a statistika – Mirko Navara
- B4B36ZUI: Základy umělé inteligence – Michal Pěchouček, Jiří Kléma
- B4B33RPZ: Rozpoznávání a strojové učení – Jiří Matas
- B4B01JAG: Jazyky, automaty, gramatiky – Marie Demlová

Magisterská etapa:

- BE4M33SSU: Statistical Machine Learning – Boris Flach
- B4M36SMU: Symbolické strojové učení – Filip Železný, Jiří Kléma
- B4M36SAN: Statistická analýza dat – Jiří Kléma



Děkuji za pozornost.

Dotazy?