

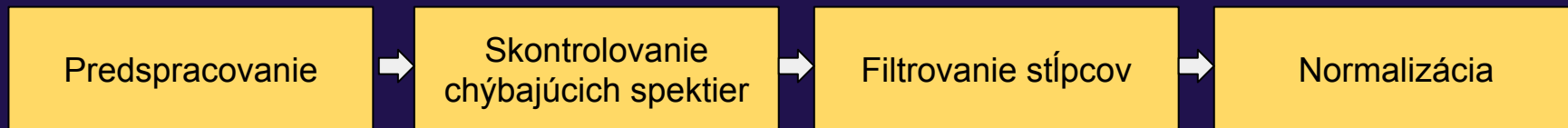
# Prezentácia produktu

# Program

- Prezentácia klasifikácie
  - Spracovanie dát
  - Vypočítanie črt
  - Trénovanie modelov
  - Vyhodnotenie výsledkov
- Web
- Aplikácia

# Spracovanie dát

Pred samotným spracovaním rozdelenie na Train/Test sadu



# Spracovanie dát

- Prevod .txt na .csv
- Všetky údaje sú zachované pre prípadné použitie

Spectra.POSITIVE.txt – Poznámkový blok

Súbor Úpravy Formát Zobrazíť Pomocník

Spectrum 1.  
Drift field intensity: 558.521 [V/cm]  
Pressure: 994.85 [mbar]  
Temperature T1: 110.292 [C]  
Drift tube length: 11.51 [cm]  
Analysis time: 0.13 [sec]

mobility

28.83505	0.034
27.20288	0.032
25.74558	0.033
24.43648	0.034
23.25407	0.034
22.18081	0.034
21.20224	0.033
20.30637	0.034
19.48314	0.033



	A	B	C	D	E	F	G	H	I	J	
1	class	oil_id	spectrum	drift_field	pressure	temperatu	drift_tube	analysis_ti	v_0	v_1	v_
2	EVOO	200427_1,1		558.521	994.85	110.292	11.51	0.13	0.034	0.032	0.
3	EVOO	200427_1,2		558.521	994.85	110.292	11.51	0.26	0.034	0.032	0.
4	EVOO	200427_1,3		558.521	994.85	110.37	11.51	0.39	0.036	0.032	0.
5	EVOO	200427_1,4		558.521	994.85	110.37	11.51	0.52	0.035	0.033	0.
6	EVOO	200427_1,5		558.521	994.85	110.37	11.51	0.65	0.035	0.032	0.
7	EVOO	200427_1,6		558.521	994.85	110.448	11.51	0.78	0.037	0.031	0.
8	EVOO	200427_1,7		558.521	994.85	110.526	11.51	0.91	0.036	0.032	0.
9	EVOO	200427_1,8		558.521	994.85	110.448	11.51	1.04	0.035	0.032	0.
10	EVOO	200427_1,9		558.521	994.85	110.526	11.51	1.17	0.034	0.033	0.
11	EVOO	200427_1,10		558.521	994.85	110.526	11.51	1.3	0.033	0.032	0.
12	EVOO	200427_1,11		558.521	994.85	110.526	11.51	1.43	0.032	0.032	0.

Predspracovanie

Skontrolovanie  
chýbajúcich spektier

Filtrovanie stĺpcov

Normalizácia

# Spracovanie dát

- Kontrola či boli zaznamenané všetky spektrá na základe času

	A	B	C	D	E	F	G	H	I	J	
1	class	oil_id	spectrum	drift_field	pressure	temperatu	drift_tube	analysis_t	v_0	v_1	v_
2	EVOO	200427_171	1	558.521	994.85	110.292	11.51	0.13	0.034	0.032	0.0
3	EVOO	200427_172	2	558.521	994.85	110.292	11.51	0.26	0.034	0.032	0.0
4	EVOO	200427_173	3	558.521	994.85	110.37	11.51	0.39	0.036	0.032	0.0
5	EVOO	200427_174	4	558.521	994.85	110.37	11.51	0.52	0.035	0.033	0.0
6	EVOO	200427_175	5	558.521	994.85	110.37	11.51	0.65	0.035	0.032	0.0
7	EVOO	200427_176	6	558.521	994.85	110.448	11.51	0.78	0.037	0.031	0.0
8	EVOO	200427_177	7	558.521	994.85	110.526	11.51	0.91	0.036	0.032	0.0
9	EVOO	200427_178	8	558.521	994.85	110.448	11.51	1.04	0.035	0.032	0.0
10	EVOO	200427_179	9	558.521	994.85	110.526	11.51	1.17	0.034	0.033	0.0
11	EVOO	200427_1710	10	558.521	994.85	110.526	11.51	1.3	0.033	0.032	0.0
12	EVOO	200427_1711	11	558.521	994.85	110.526	11.51	1.43	0.032	0.032	0.0

Predspracovanie



Skontrolovanie  
chýbajúcich spektier



Filtrovanie stĺpcov

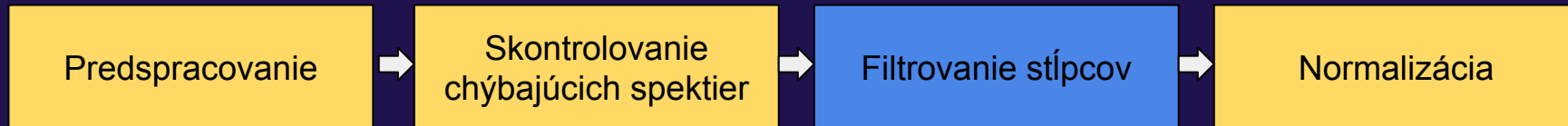
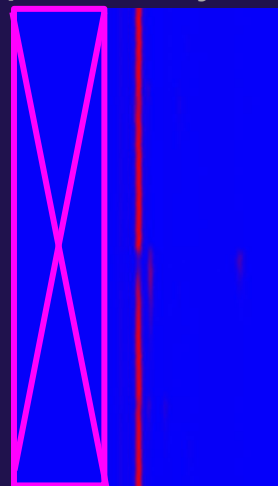


Normalizácia

# Spracovanie dát

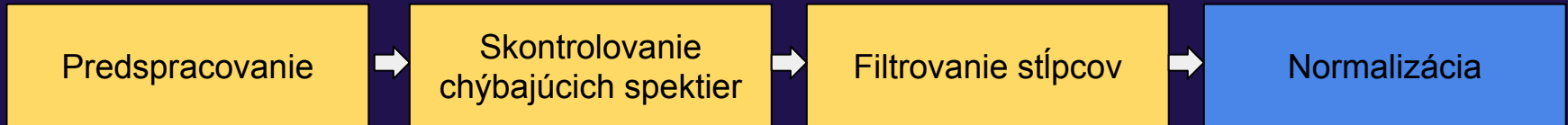
- Odstránenie stĺpcov na základe analýzy => Zmenšenie veľkosti súborov
- Odstránenie časov, prvých 150 hodnôt, poslednej hodnoty

```
Spectra.POSITIVE.txt - Poznámkový blok
Súbor Úpravy Formát Zobrazit Pomocník
Spectrum 1.
Drift field intensity: 558.521 [V/cm]
Pressure: 994.85 [mbar]
Temperature T1: 110.292 [C]
Drift tube length: 11.51 [cm]
Analysis time: 0.13 [sec]
mobility
28.83505 0.034
27.20288 0.032
25.74558 0.033
24.43648 0.034
23.25407 0.034
22.18081 0.034
21.20224 0.033
20.30637 0.034
19.48314 0.033
```



# Spracovanie dát

- 0-1 normalizácia
- všetky hodnoty sú v rozsahu 0-1



# Vypočítanie črt

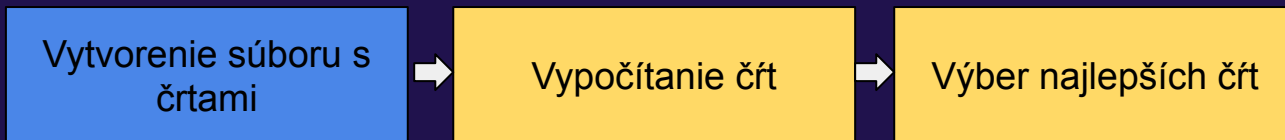




# Vypočítanie črt

- Pri vytváraní sa vypočítajú črty
- Ak súbor existuje, doplnia sa len chýbajúce črty => ľahšie pridávanie nových

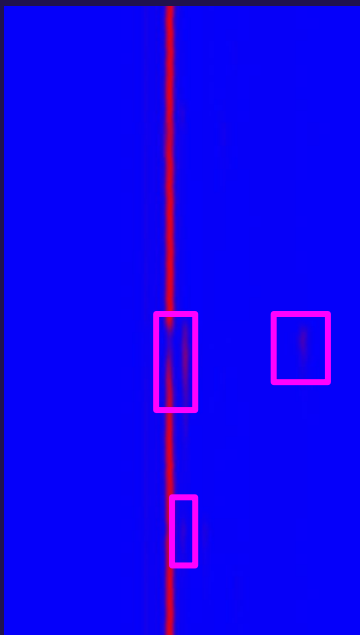
	A	B	C	D	E	F	G
1	class	oil_id	std_v_287	std_v_288	std_v_289	std_v_290	std_v_291
2	EVOO	200427_190222	0.0023507	0.0022104	0.0020219	0.0019874	0.001769
3	EVOO	200428_011038	0.0027605	0.0027966	0.0024850	0.0023687	0.002003
4	EVOO	200428_215357	0.0024819	0.0022505	0.0023167	0.0020253	0.001994
5	EVOO	200428_230730	0.0022985	0.0021052	0.0020654	0.0020376	0.001820
6	EVOO	200429_160212	0.0023721	0.0023588	0.0024931	0.0021252	0.002045
7	EVOO	200429_171550	0.0025380	0.0023763	0.0023527	0.0024115	0.002675
8	EVOO	200429_182924	0.0023257	0.0021199	0.0020594	0.0019913	0.002040
9	EVOO	200430_064535	0.0029600	0.0028402	0.0025864	0.0022661	0.001995
10	EVOO	200430_125454	0.0030120	0.0029931	0.0029306	0.0028354	0.002555



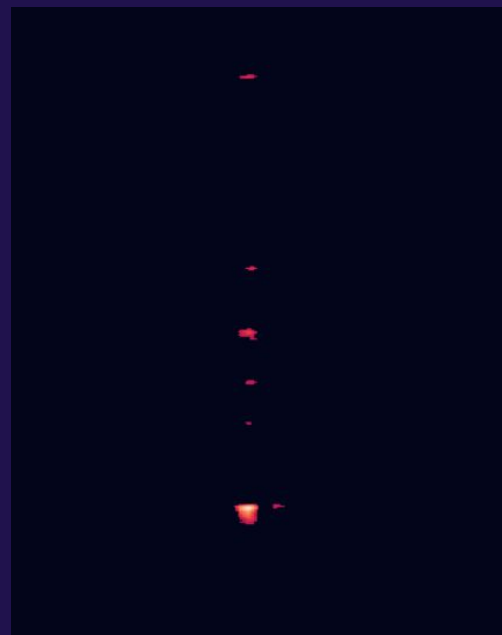
# Vypočítanie črt

- mean
- max
- min
- median
- std
- delta (max-min)
- sum
  
- diff (mask)
- point (max/mean)

Point



diff



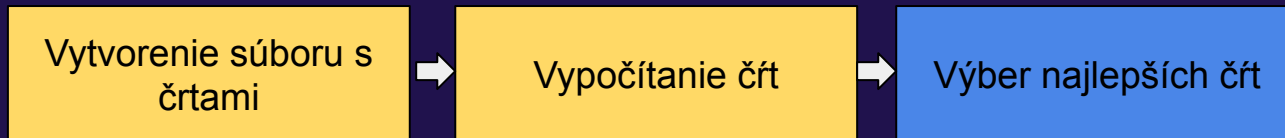
Vytvorenie súboru s  
črtami

Vypočítanie črt

Výber najlepších črt

# Vypočítanie črt

- Použité na črty: mean, max, min, median, std, delta, sum
- 7 črt x 298 stĺpcov = 2086 črt
- Výber = 102 črt



# Trénovanie modelov

- Random Forest
- Decision Tree
- Knn
- Naive Bayes
- SVM
- MLP

# Vylepšenia modelov cez hyperparametre

- GridSearch
  - skúšanie všetkých kombinácií vybraných parametrov

## Príklad Random Forest:

```
tuned_parameters = {  
    'n_estimators': [int(x) for x in np.linspace(start=10, stop=160, num=20)],  
    'criterion': ['gini', 'entropy'],  
    'min_samples_split': [int(x) for x in np.linspace(start=2, stop=30, num=2)],  
    'bootstrap': [True, False],  
    'max_depth': [None, 5, 10, 20],  
    'max_leaf_nodes': [None, 5, 10, 15],  
    'max_samples': [None, 5, 10, 15]  
}
```

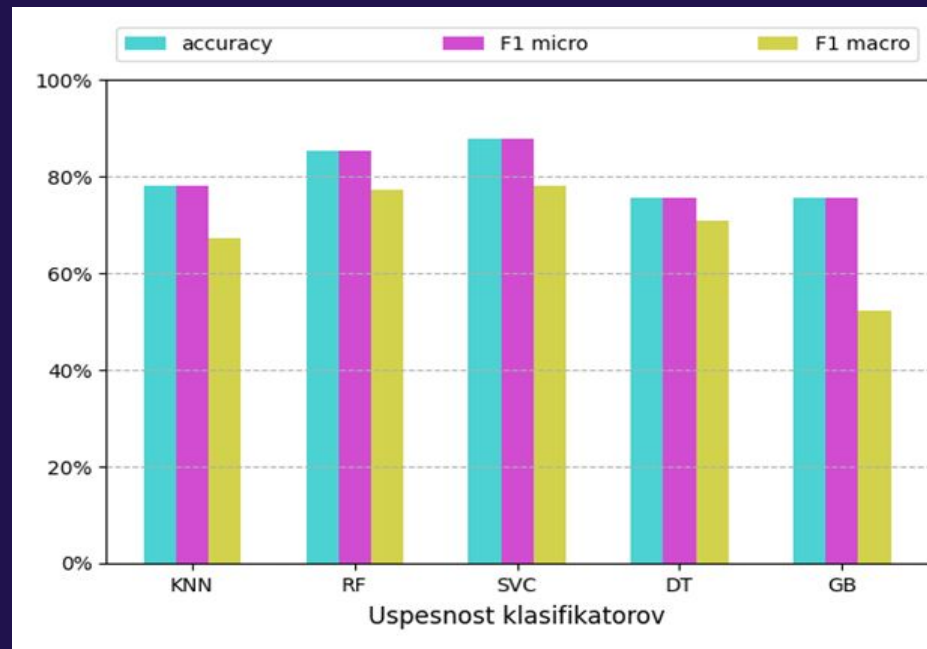
# Úspěšnost' modelov

	Accuracy	Fáza (použité črty)
1.	<b>51%-65% (raz 78%)</b>	<b>mean_ALL, max v_0</b>
2.	77%-88%	max, mean
3.	52%-78% / 56%-73%	point / diff mapy
4.	61%-76%	výber čřt

```
Connected to pydev debugger (bu
accuracy 0.5365853658536586
macro 0.46198482368695143
micro 0.5365853658536586
[[14  3  3]
 [ 5  7  0]
 [ 8  0  1]]
```

# Úspěšnost' modelov

	Accuracy	Fáza (použité črty)
1.	51%-65% (raz 78%)	mean_ALL, max v_0
<b>2.</b>	<b>77%-88%</b>	<b>max, mean</b>
3.	52%-78% / 56%-73%	point / diff mapy
4.	61%-76%	výber čřt



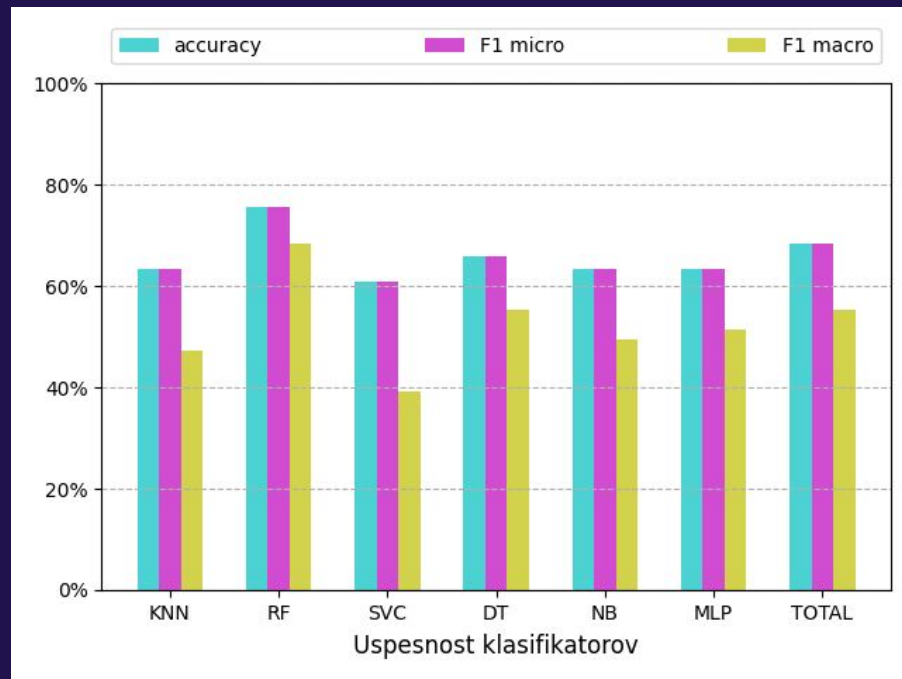
# Úspěšnost' modelov

	Accuracy	Fáza (použité črty)
1.	51%-65% (raz 78%)	mean_ALL, max v_0
2.	77%-88%	max, mean
<b>3.</b>	<b>52%-78% / 56%-73%</b>	<b>point / diff</b>
4.	61%-76%	výber črt



# Úspěšnost' modelov

	Accuracy	Fáza (použité črty)
1.	51%-65% (raz 78%)	mean_ALL, max v_0
2.	77%-88%	max, mean
3.	52%-78% / 56%-73%	point / diff mapy
4.	<b>61%-76%</b>	<b>výber črt</b>



# Web

- pripravené a implementované obrazovky
- implementovaný backend
- podpora pre takmer všetky veľkosti zariadenia
- postupná údržba

# Aplikácia

- pripravené mock-upy
- implementovaná prvá obrazovka