

Vilniaus Universitetas

Matematikos ir Informatikos fakultetas

Daugiamačių duomenų vizualizavimas

2 Užduotis

Tiesioginio vizualizavimo metodai

Laimonas Beniušis
Informatika 1gr.

Turiny

1	Visualizavimo metodai.....	3
2	Atributai.....	3
3	Duomenų vizualizavimas.....	4
3.1	Attribute graph (atributų grafikai).....	4
3.2	Scatter Plot.....	5
3.3	Andrew Curves (Andrew kreivės).....	6
3.4	FreeViz.....	7
3.5	RadViz.....	8
3.6	Parallel Coordinates (lygiagrečių koordinačių vizualizacija).....	9

1 Visualizavimo metodai

- *Attribute Graph*
- *Scatter Plot*
- *Andrew Curves*
- *FreeViz*
- *RadViz*
- *Parallel Coordinates*

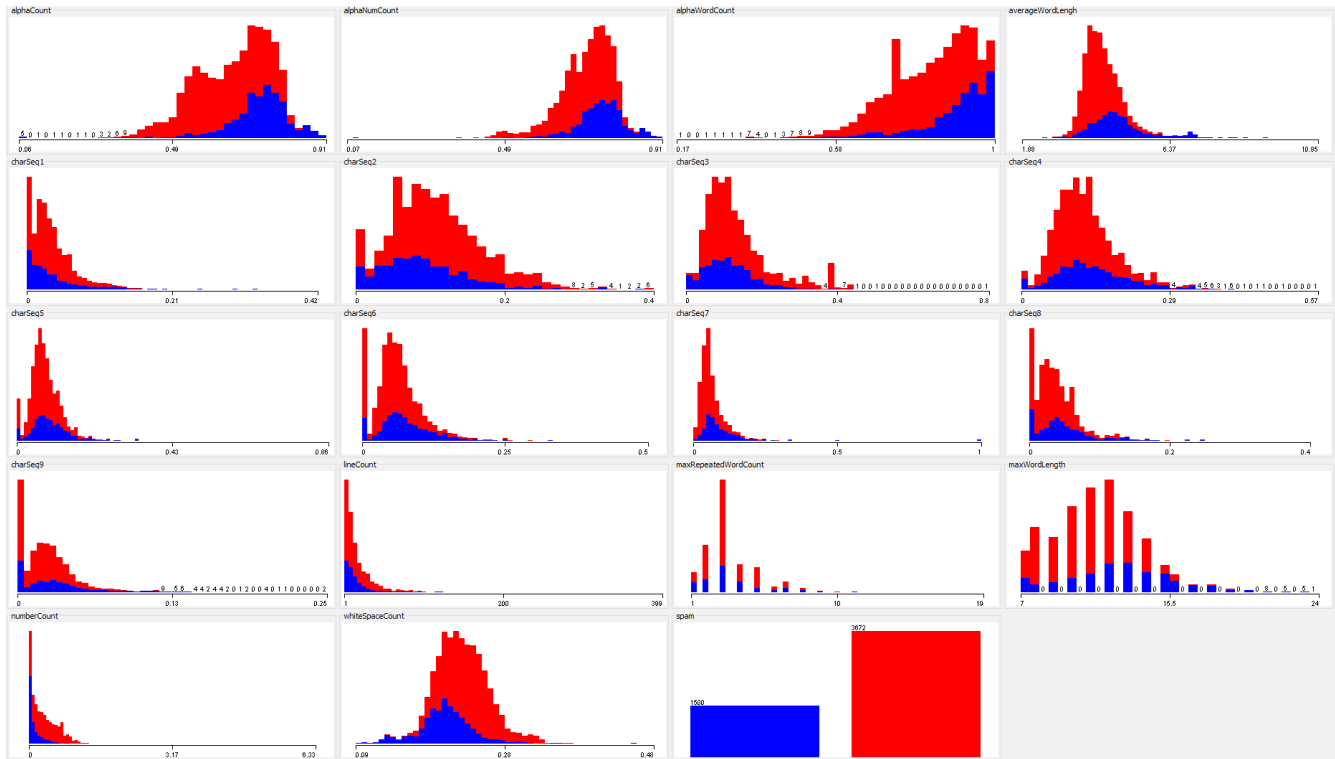
2 Atributai

Duomen aibės atributai yra ženkliai sumažinti iki \sim (770 ir 150) (buvo 5172). Ne visi atributai yra tinkami vizualizavimui, todėl palikti tik šie:

alphaCount	Raidžių kiekis / simbolių kiekis
alphaNumCount	Raidžių ar skaičių kiekis/ simbolių kiekis
alphaWordCount	Raidinių žodžių kiekis / žodžių kiekis
charSeq1	Raidinių žodžių kiekis, kurių ilgis 1 / žodžių kiekis
charSeq2	Raidinių žodžių kiekis, kurių ilgis 2 / žodžių kiekis
charSeq3	Raidinių žodžių kiekis, kurių ilgis 3 / žodžių kiekis
charSeq4	Raidinių žodžių kiekis, kurių ilgis 4 / žodžių kiekis
charSeq5	Raidinių žodžių kiekis, kurių ilgis 5 / žodžių kiekis
charSeq6	Raidinių žodžių kiekis, kurių ilgis 6 / žodžių kiekis
charSeq7	Raidinių žodžių kiekis, kurių ilgis 7 / žodžių kiekis
charSeq8	Raidinių žodžių kiekis, kurių ilgis 8 / žodžių kiekis
charSeq9	Raidinių žodžių kiekis, kurių ilgis 9 / žodžių kiekis
numberCount	Skaičių (žodžių iš skaičių) kiekis / žodžių kiekis
whiteSpaceCount	Tarpinių simbolių (<i>whitespace</i>) kiekis / simbolių kiekis
spam	{SPAM,NOTSPAM} E.laiško kategorija

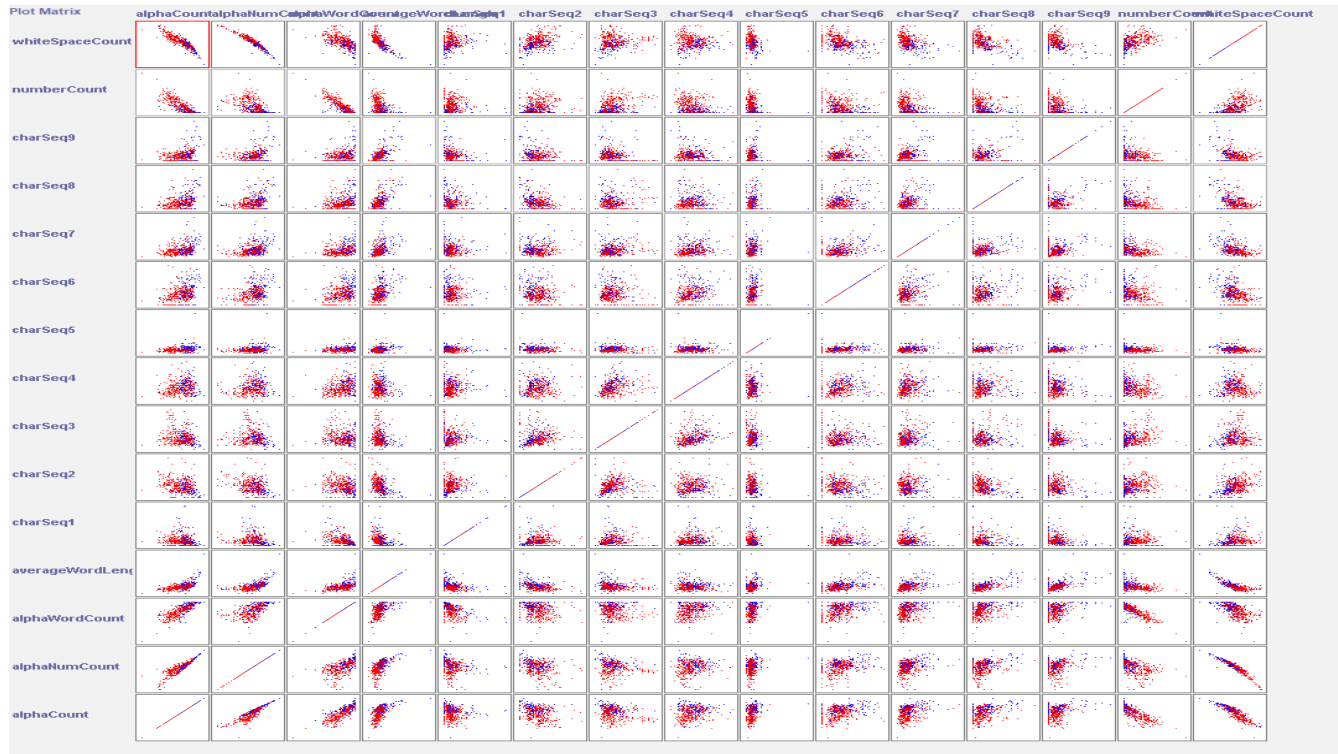
3 Duomenų vizualizavimas

3.1 Attribute graph (atributų grafikai)



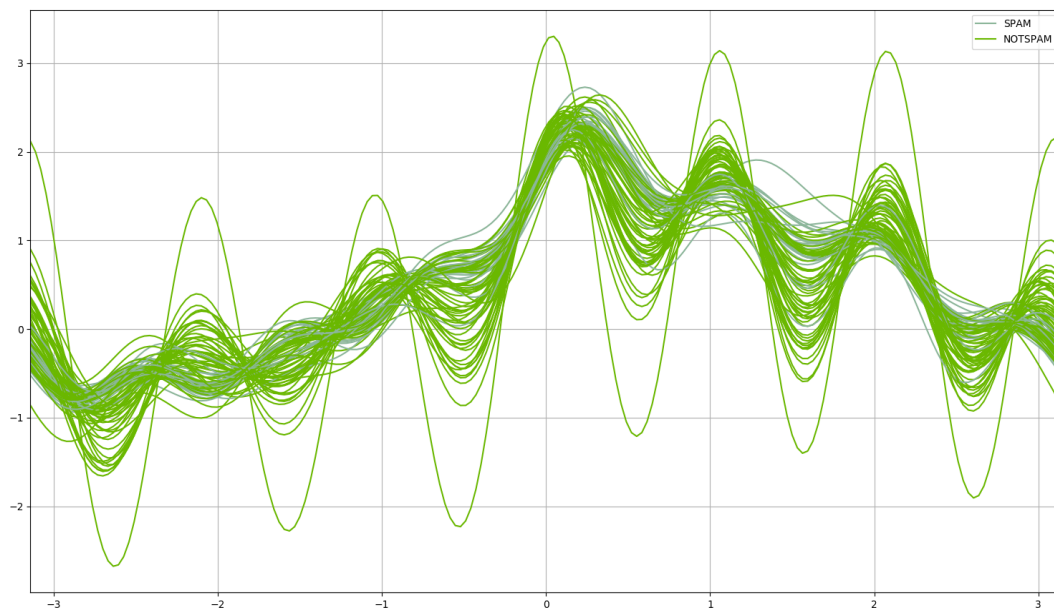
1. Pav: Weka duomenų atributų grafikai

3.2 Scatter Plot

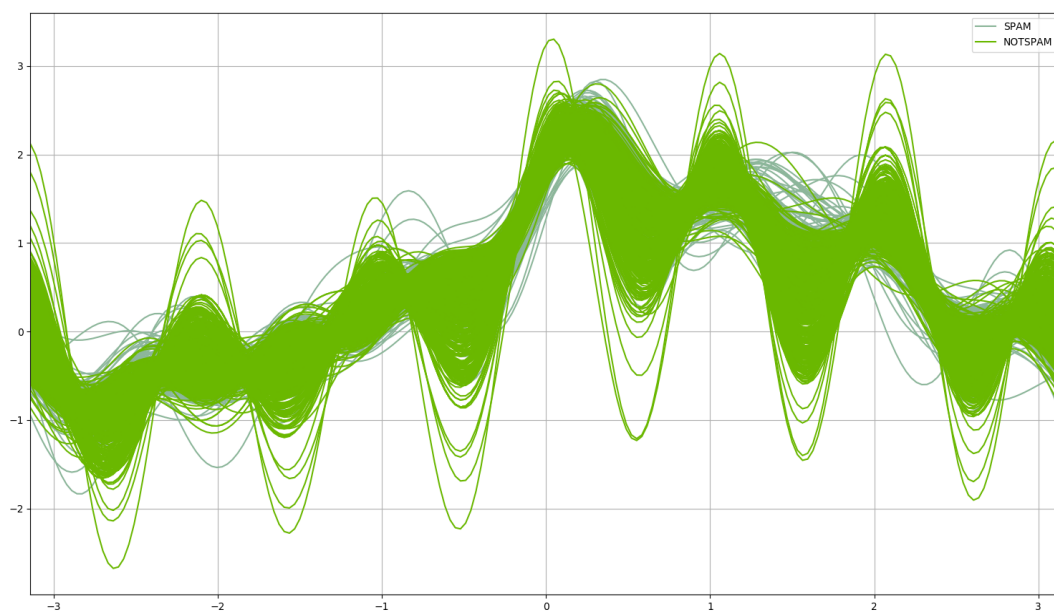


2. Pav: Weka duomenų Scatter Plot matrica

3.3 Andrew Curves (Andrew kreivės)

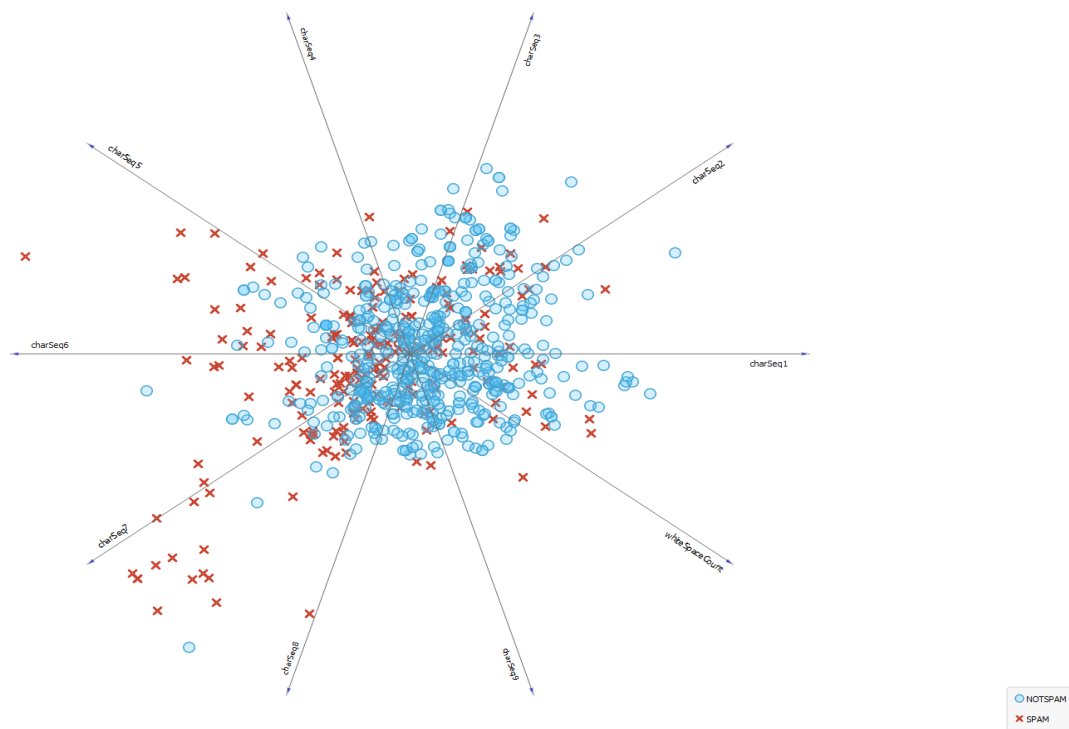


3. Pav: Pandas duomenų Andrew kreivės, 155 įrašų

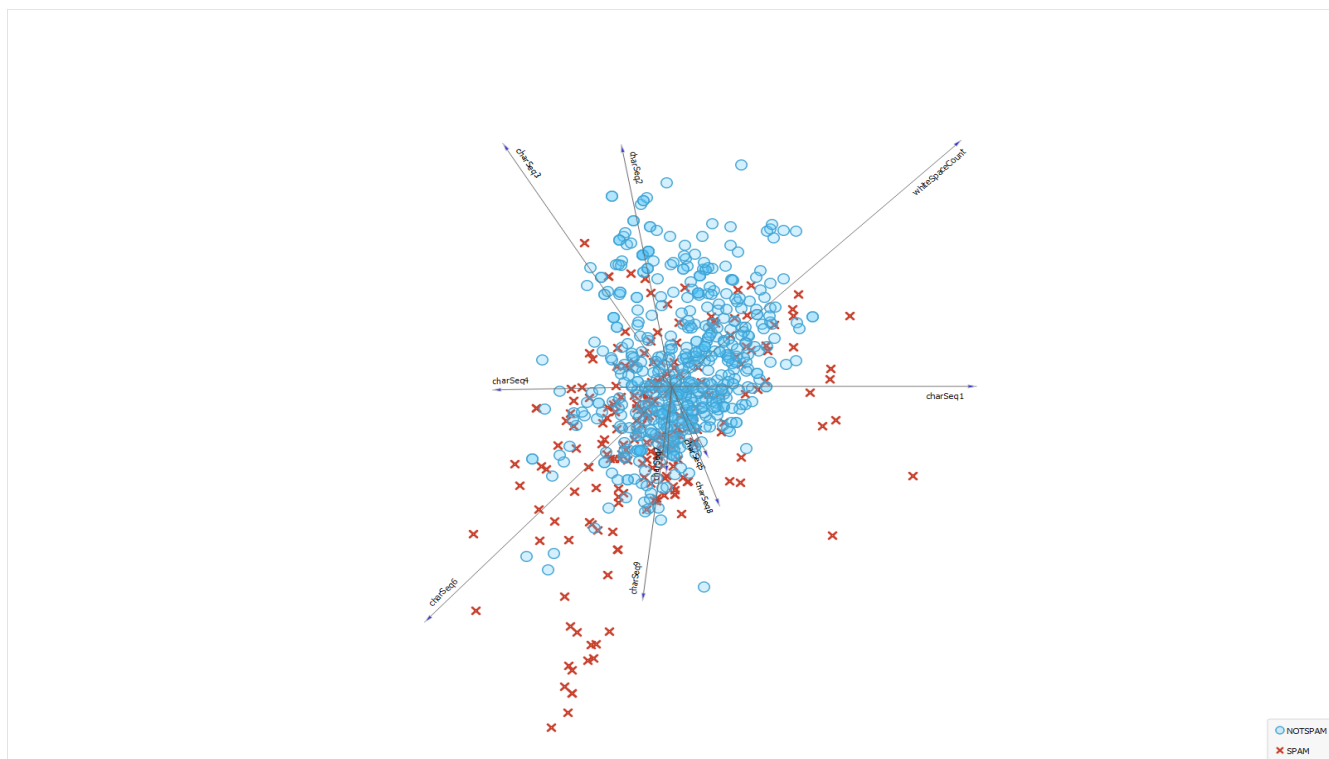


4. Pav: Pandas duomenų Andrew kreivės, 770 įrašų

3.4 FreeViz

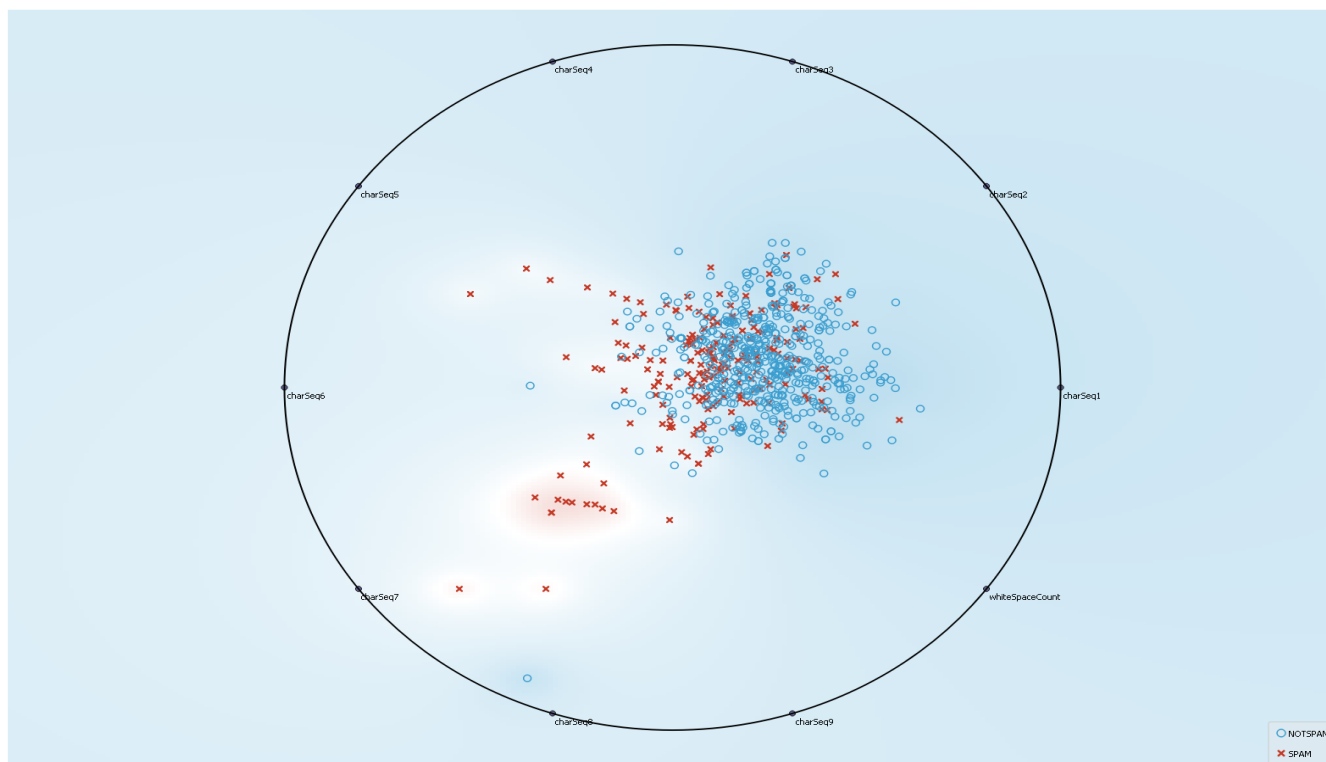


5. Pav: Orange duomenų FreeWiz vizualizacija (neoptimizuota), 770 įrašų



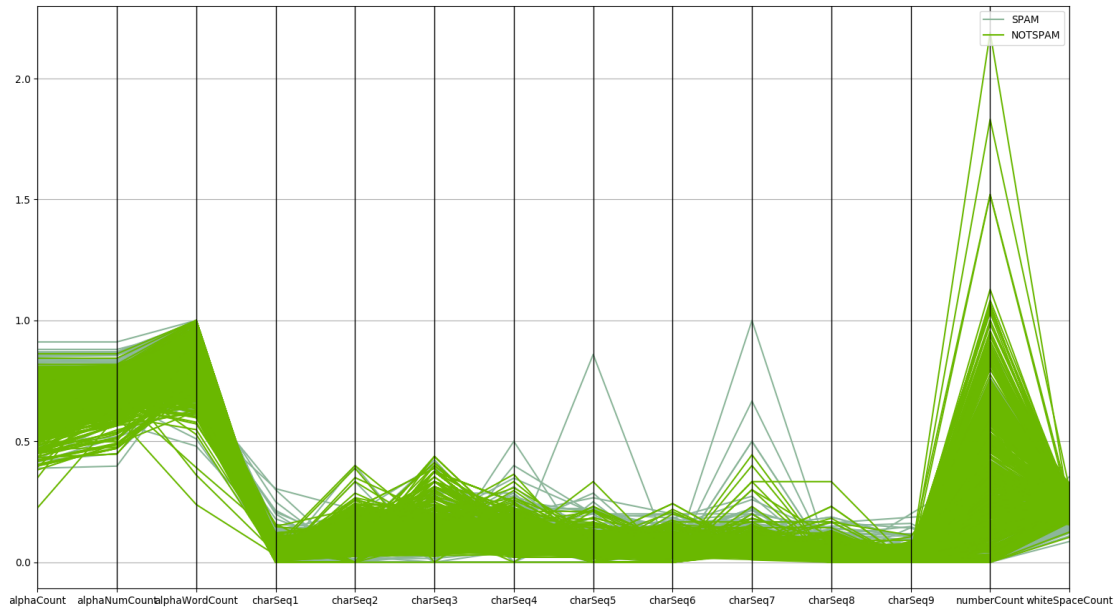
6. Pav: Orange duomenų FreeWiz vizualizacija (optimizuota), 770 įrašų

3.5 RadViz

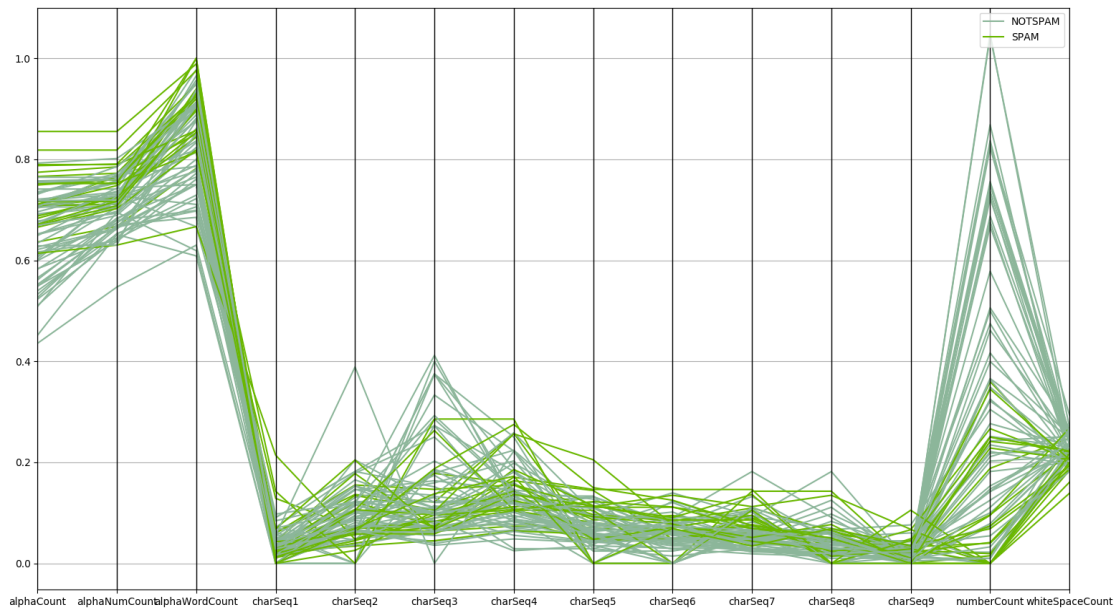


7. Pav: Orange duomenų RadViz vizualizacija , 770 įrašų

3.6 Parallel Coordinates (lygiagrečių koordinačių vizualizacija)



8. Pav: Pandas duomenų lygiagrečių koordinačių vizualizacija , 770 įrašų



9. Pav: Pandas duomenų lygiagrečių koordinačių vizualizacija , 150 įrašų