

Визуелизација

Ненад Митић

Математички факултет
nenad@matf.bg.ac.rs

Uvod

Визуелизација је конверзија података у визуелни или табеларни формат

- Анализа и формирање извештаја о карактеристикама података и односима у њима
- Могућност анализе визуелног приказа велике количине података
- Откривање општих образаца и трендова
- Откривање елемената ван граница и нетипичних образаца

Избор врсте визуелизације

- Начин и врста приказа зависе од података
- Елиминација/смањење наглашености појединих атрибута или подскупа атрибута
- Димензиона редукција ради смањења броја димензија на две или три
- Алтернативно, посматра се више парова атрибута

Избор врсте визуелизације (наставак)

- Избор може да укључи издвајање скупа објеката
- Део екрана садржи превише тачака → проређивање
- У случају проређивања, тачке у проређеним деловима не треба дирати

Primeri:

- Glenn Myatt and Wayne Johnson: Making Sense of Data II: A Practical Guide to Data Visualization, Advanced Data Mining Methods, and Applications, John Wiley and Sons, 2009
- Danyel Fisher and Miriah Meyer: Making Data Visual - A Practical Guide to Using Visualization for Insight, O'Reilly 2018
- Nathan Yau: Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics, John Wiley and Sons, 2011

Пресликавање података у графичке елементе

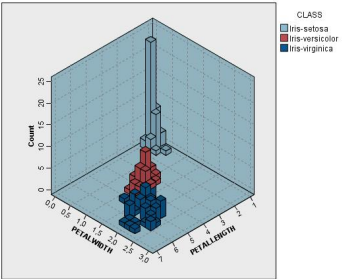
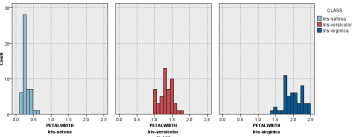
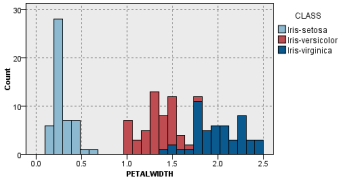
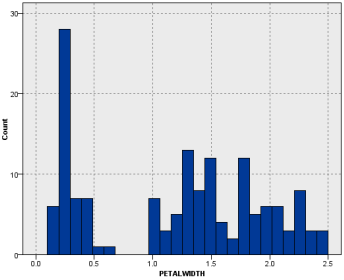
- Објекти, атрибути и њихови односи се преводе у графичке елементе као што су тачке, линије, облици, боје,
- Пример:
 - Објекти се често приказују као тачке
 - Вредности атрибута могу да се прикажу као позиција тачке или као њене карактеристике (боја, величина, облик,...)
 - Ако се за представљање користи позиција лако се уочавају односи (групе, облик, елементи ван граница, ...)

Уређеност

Начин смештања визуелних елемената може јако да утиче на разумевање

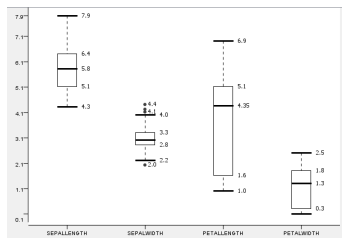
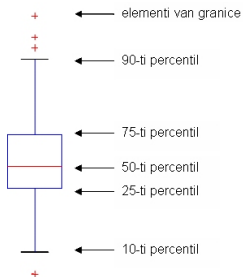
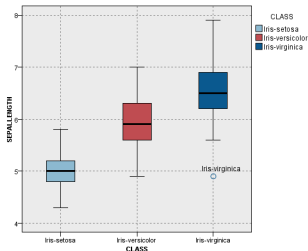
	1	2	3	4	5	6
1	0	1	0	1	1	0
2	1	0	1	0	0	1
3	0	1	0	1	1	0
4	1	0	1	0	0	1
5	0	1	0	1	1	0
6	1	0	1	0	0	1
7	0	1	0	1	1	0
8	1	0	1	0	0	1
9	0	1	0	1	1	0

Хистограми



Кућице

- Кућице (eng. *box plot* - установио их је *J. Tukey*)
- Dodatni начин за приказивање дистрибуције података



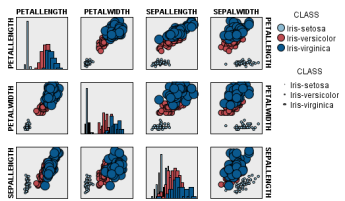
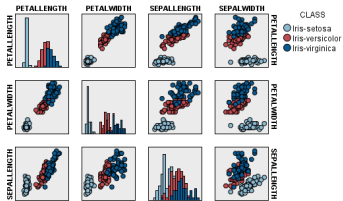
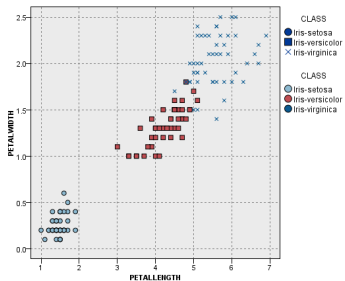
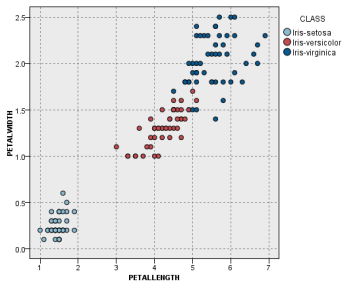
Перцентили

Ниво (линије) у кућицама су дати у перцентилима. За нумерички атрибут x и број $p \in [0, 100]$, p -ти перцентил је вредност x -а таква да је $p\%$ вредности x мање од ње. Нпр. 50-ти перцентил је вредност од које је мање 50% вредности у x -у

Шеме са распршеним елементима

- Атрибути вредности одређују позицију
- Најчешће се користе *2D* али постоје и *3D* шеме
- Додатни атрибути могу да се прикажу коришћењем облика, величине или боје маркера који представљају објекте
- Корисно је имати низове шема које приказују сумарне односе више парова атрибута

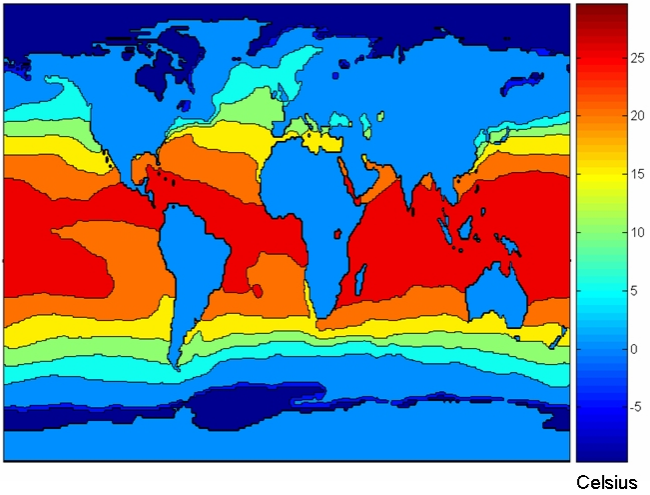
Шеме са распршеним елементима



Шеме са контурама

- Корисне су када се непрекидни атрибути мере на просторној основи
- Деле раван у регионе са сличним вредностима
- Контурне линије које чине границе региона повезују тачке са једнаким вредностима
- Најчешћи пример је контурна мапа са висинама (нпр. терен са изохипсама)
- Приказ температуре, падавина, ваздушног притиска,
...

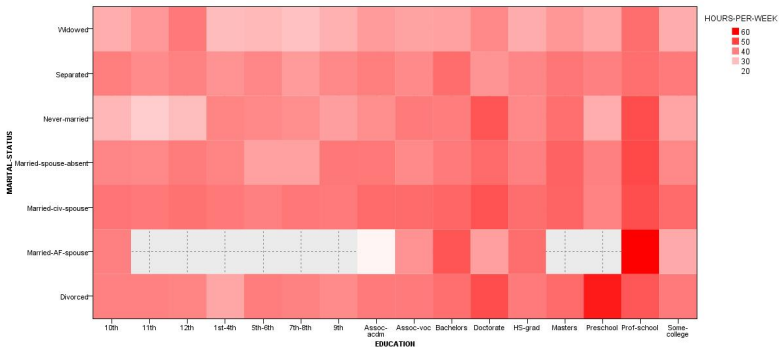
Шеме са контурама



Шеме са матрицама

- Чест назив - *Топлотне мапе*
- Корисне су када су објекти сортирани у класе
- Уобичајено је да се атрибути нормализују да би се избегла доминација једног атрибута над другим
- Шеме са матрицама које садрже сличности или разлике могу да буду јако корисне у представљању односа међу објектима

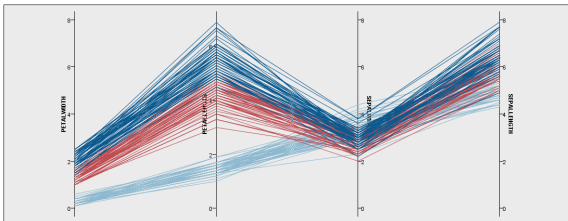
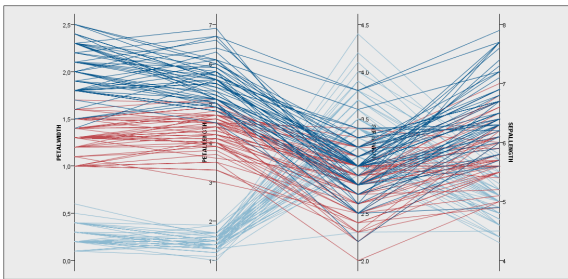
Шеме са матрицама



Паралелне координате

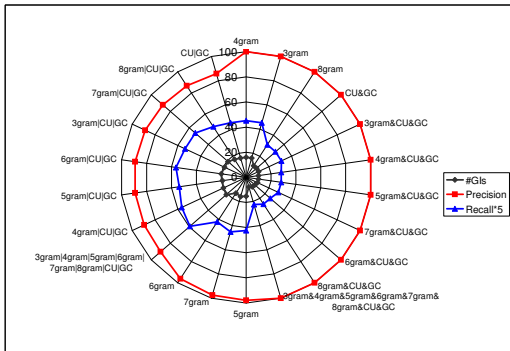
- Користе се за приказ вредности атрибута вишедимензионалних података
- Уместо управних оса користе се паралелне
- Сваки објекат се представља линијом тако што се вредности његових атрибута представљају тачкама на свакој одговарајућој координатној оси, при чему се тачке повезују линијом
- За неке атрибуте линија често приказује различите класе објеката који су груписани заједно
- Уређење атрибута је важно за уочавање груписања

Паралелне координате

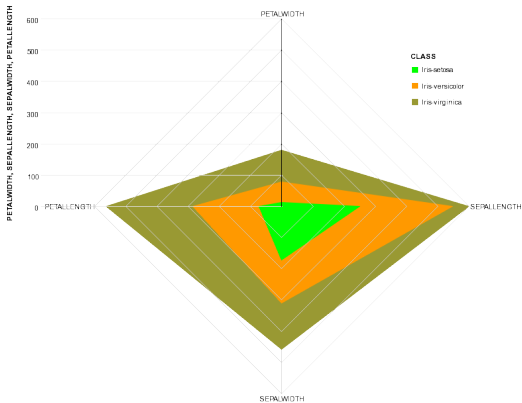


Цртање звезда

- Приступ је сличан као код паралелних координата, али се осе гранају из једне тачке у центру
- Тачке објеката су спојене полигоналном линијом



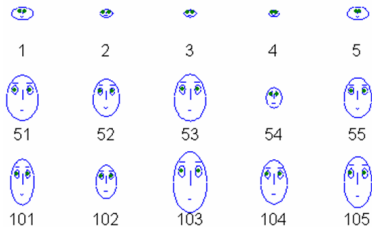
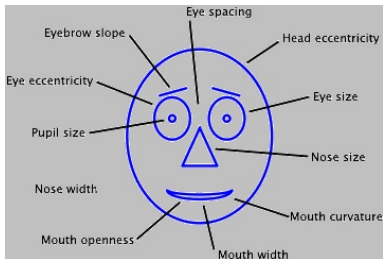
Цртање звезда



Chernoff-ова лица

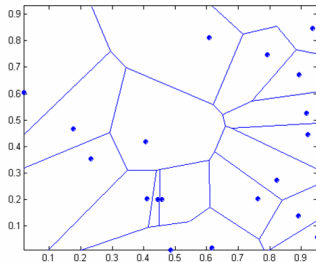
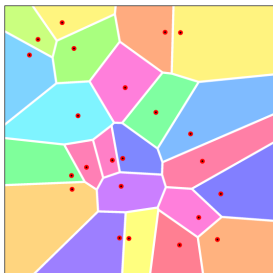
- Приступ је формирао *Herman Chernoff*
- Приступ се заснива на разликовању особина лица од стране посматрача
- Сваки атрибут се придружује некој карактеристици лица
- Вредност атрибута одређује изглед и карактеристике лица
- Објекат се представља као појединачно лице

Chernoff-ова лица (наставак)



Дијаграми *Voronoi*-а

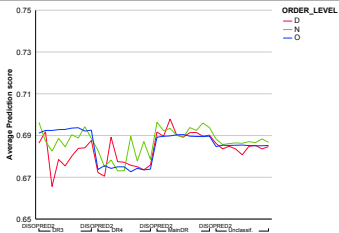
За дати скуп тачака дијаграм *Voronoi*-а је подела простора у регионе унутар којих су све тачке ближе неком појединачном чвору него било ком другом чвору



Слике и табеле

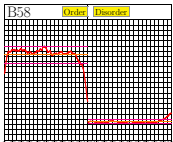
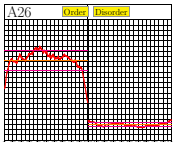
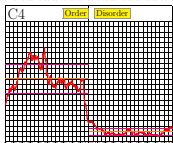
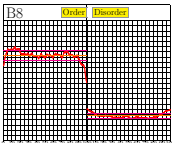
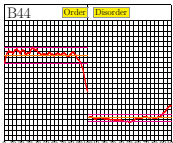
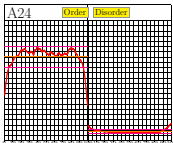
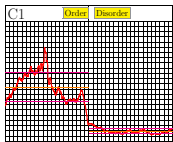
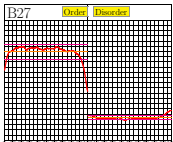
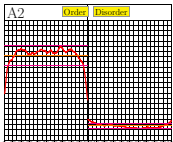
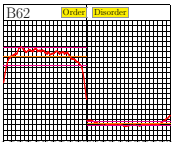
CLASSIFICATION: Greenbaum

BIND_LEVEL: SB

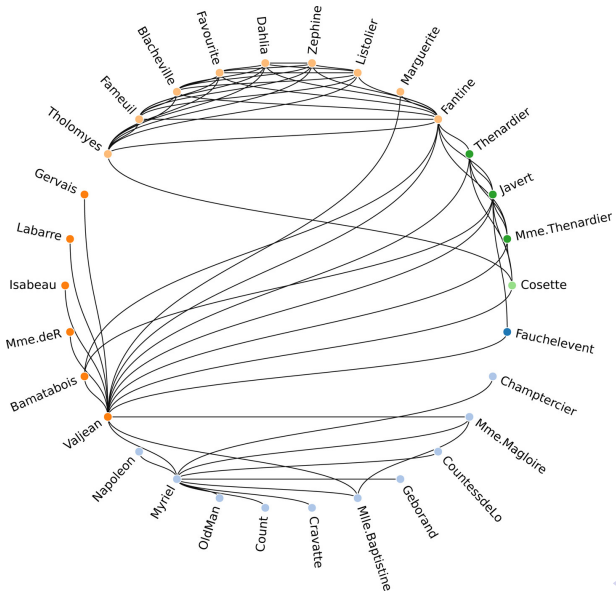


		SUPERTYPE_PREDICTOR								
Average Prediction score		DISOPRED2	DisEMBL_Hot_loops	DisEMBL_Remark465b	IUPred-L	IUPred-S	IsUnstruct	OnDCRF	RONN	VSL2b
DR3	D	0.686	0.692	0.666	0.679	0.675	0.680	0.684	0.684	0.688
	N	0.696	0.687	0.682	0.689	0.685	0.690	0.689	0.694	0.689
	O	0.691	0.692	0.692	0.693	0.693	0.694	0.694	0.692	0.693
DR4	D	0.672	0.670	0.689	0.678	0.677	0.676	0.675	0.674	0.676
	N	0.683	0.675	0.678	0.673	0.673	0.690	0.678	0.687	0.679
	O	0.674	0.676	0.674	0.675	0.675	0.673	0.674	0.674	0.674
MainDR	D	0.692	0.690	0.698	0.690	0.689	0.691	0.691	0.690	0.690
	N	0.696	0.692	0.693	0.690	0.689	0.694	0.692	0.696	0.694
	O	0.689	0.690	0.690	0.690	0.690	0.690	0.690	0.689	0.690
Unclassif.	D	0.686	0.684	0.685	0.684	0.681	0.685	0.685	0.684	0.685
	N	0.688	0.686	0.686	0.686	0.686	0.687	0.687	0.688	0.687
	O	0.685	0.685	0.685	0.685	0.685	0.685	0.685	0.685	0.685

Поређење



Чворови у мрежи (наставак)



Мрежа повезаности

