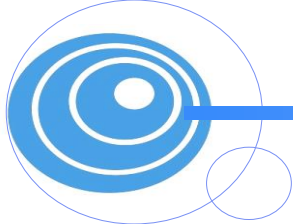


Наука о подацима у саобраћају

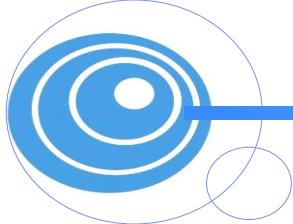
Наставници (*e-mail*):

- др Ана Узелац, доцент
(ana.uzelac@sf.bg.ac.rs)
- др Катарина Кукић, ванредни професор
(k.mijailovic@sf.bg.ac.rs)
- др Слађана Јанковић, ванредни професор
(s.jankovic@sf.bg.ac.rs)



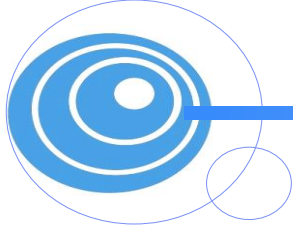
Садржај

- Наука о подацима (Data science) – увод.
- Различити типови података.
- Припрема података. Рад са библиотекама у Python-у или неком сличном окружењу које служе за процесирање података.
- Важнији појмови из области Линеарне алгебре и Вероватноће и статистике потребни за рад са подацима.
- Технике визуелизације података. Рад са библиотекама у *Python*-у или неком сличном окружењу које служе за визуелизацију података.



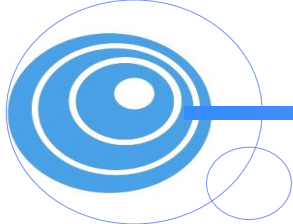
Садржај

- Алгоритми машинског учења.
- Предиктивна анализа која се базира на моделима машинског учења (алгоритми класификације, регресије и кластеровања).
- Грешке и методе за евалуацију модела машинског учења.
- Примери примене data science-а у решавању проблема у саобраћају.



Циљ предмета

- Студент се припрема за рад са скуповима великих података укључујући прикупљање, истраживање и моделирање података. Ради се на развијању скупа вештина потребних за послове који подразумевају знање из примењене математике, информатике, аналитике података и статистичких метода.
- Упознавање студената са основним концептима, техникама и алатима за решавање:
 - класификационих и регресионих задатака,
 - задатака кластеровања у саобраћају, транспорту или логистици,
 - евалуацију тачности модела и визуелизација добијених решења.

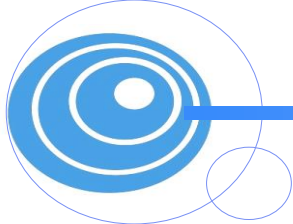


Циљ предмета

Студенти ће се упознати са:

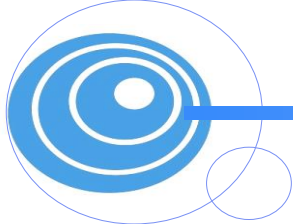
- различитим типовима података,
- бар једним софтверским алатом који се користи за обраду података,
- основним алгоритмима машинског учења, прављењем, валидацијом и тестирањем предиктивних модела, визуелизацијом података.

Главни циљ је примена усвојених техника data science-а на анализу и решавање проблема из саобраћаја.

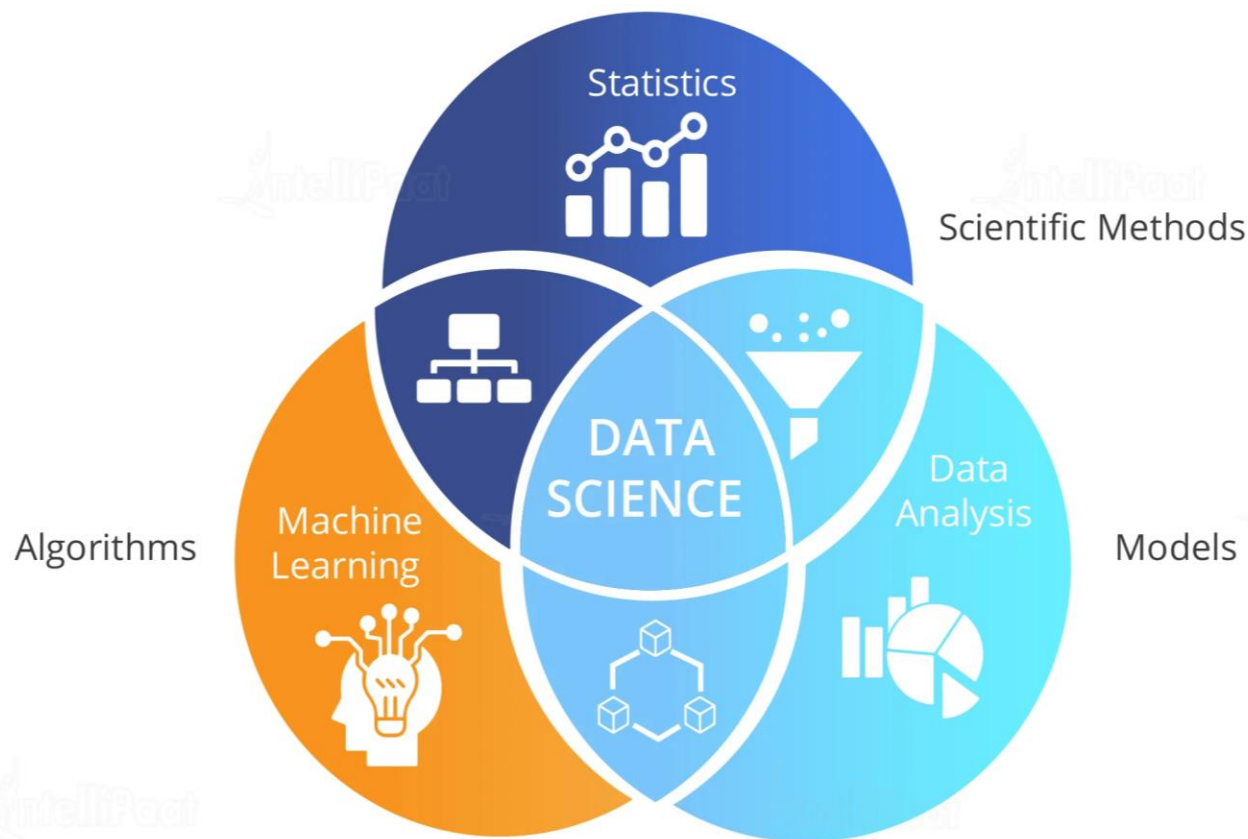


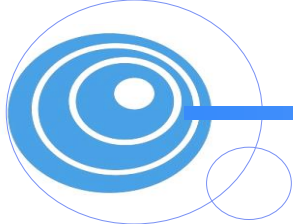
Појам Data Science

- Под појмом Data Science или Наука о подацима подразумева се мултидисциплинарна област која користи анализу података из којих помоћу математичких и статистичких метода и машинског учења извлачи закључке и на основу којих даје предикције.
- Са наглим развојем информационих технологија и урбанизације, постала је доступна велика количина података из различитих извора и различите природе. Питање како искористити те количине података је једно од најактуелнијих питања данас.
- Може се рећи да смо ушли у нову еру доношења одлука као резултат повећања способности складиштења, приступа и анализе података у многим областима, па тако и у саобраћају.



Наука о подацима (Data science)

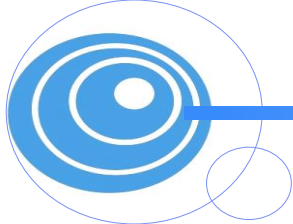




Наука о подацима (Data science)

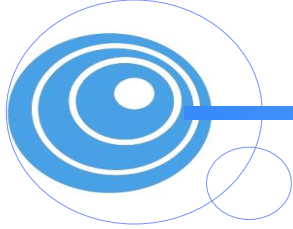
Уз помоћ науке о подацима може се обавити дескриптивна, дијагностичка, предиктивна и прескриптивна анализа, односно добити увид и предвиђања, као и одговори на питања:

- Шта се догодило?
- Зашто се то догодило?
- Шта се могло догодити?
- Шта треба учинити?



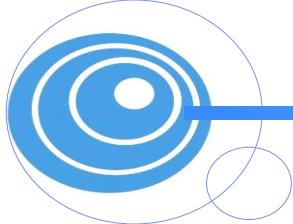
Технике и алати које ћемо изучавати у оквиру курса:

- **Data engineering:** задаци чишћења података (узимање релевантних информација) и спајања (мешање података различитог порекла).
- **Visualization:** Визуелизација података, инфографике.
- **Advanced Computing:** Имплементација у разним софтверским пакетима у зависности од предзнања студената (Python, R, Matlab).



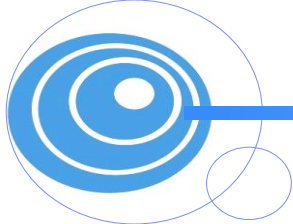
Data science - примена у саобраћају

- Аутономна возила
- Контрола саобраћаја са циљем смањења загушења
- Побољшање безбедности на путевима
- Оптимизација потрошње горива авиона
- Употреба биометријске технологије и посебно опције препознавања лица на аеродромима
- Одређивање преференција путника на основу архивираних података



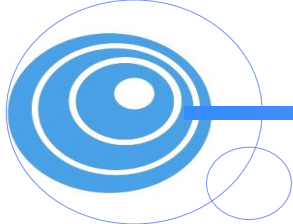
Data science у саобраћају – постојећи примери из праксе

- **StreetLight** – комбинује машинско учење, податке из мобилних телефона и знање експерата у моделовању саобраћајних образаца везаних за аутомобиле, бицикле и пешаке, како би заменили дуготрајно прикупљање података и ручно пребројавање саобраћаја.
- **Network Planning Tools (NPT)** – уз помоћ машинског учења и вештачке интелигенције оптимизује се транспорт пакета од врата до врата у компанији UPS (*American Messenger Company* in Seattle, Washington). За решавање изазовних логистичких задатака, као што је на пример преусмеравање пакета у случају лошег времена. Овај алат омогућава инжењерима да симулирају различита решења и изаберу најбоље, као и да предлаже руте.
- **Uber Eats** – апликација за доставу хране. Користи машинско учење и напредне статистичке методе са циљем да храну достави топлу без обзира на временске услове, саобраћајна загушења и слично.



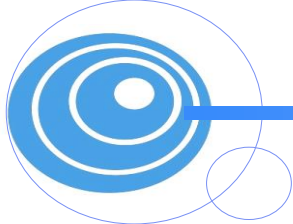
Провера знања

- **Пројектни задатак (обавезан)** – самосталан рад студента уз консултације са предметним наставницима: **40 бодова**.
- **Семинар (обавезан)** – семинарски рад или публиковани научно-стручни рад студента у сарадњи са предметним наставницима: **30 бодова**.
- **Усмени испит (није обавезан)** - може бити у форми излагања научно-стручног рада студента у сарадњи са предметним наставницима на конференцији: **30 бодова**.



Data science у саобраћају – извесна будућност





Литература

1. Sinan Ozdemir: *Principles of Data Science*, Packt Publishing Ltd, Birmingham, 2016, p. 363
2. Matlab documentation: „Introducing Machine learning“, <https://www.mathworks.com/help/stats/supervised-learning-machine-learning-workflow-and-algorithms.html#bswluhd>
3. Joel Grus, *Data Science from Scratch: First Principles with Python 2nd Edition*, O'Reilly, 2019.
4. Ian Witten, Eibe Frank, Mark Hall, Christopher Pal, *Data Mining: Practical Machine Learning Tools and Techniques, 4th edition*, Morgan Kaufmann, 2016.
5. Sridhar Alla, *Big Data Analytics with Hadoop 3*, Packt Publishing Ltd, 2018.
6. Doug Bierer, *MongoDB 4 Quick Start Guide*, Packt Publishing Ltd, 2018.