**Табела 5.2.** Спецификација предмета

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Студијски програм: Климатске промене и адаптација на климатске промене** | | | | |
| **Назив предмета: Увод у науку о подацима** | | | | |
| **Наставник: др Ненад Королија** | | | | |
| **Статус предмета: изборни** | | | | |
| **Број ЕСПБ:** 5 | | | | |
| **Услов:** Програмирање | | | | |
| **Циљ предмета**  Основни циљ овог предмета је оспособљавање студената у коришћењу програмског језика Пајтон у припреми, обради и анализи података с нагласком на употребу библиотека Numpy и Pandas, те визуализацији података коришћењем библиотеке Matplotlib. Поред основних појмова дескриптивне статистике, студенти ће бити упознати са основним надгледаним и ненадгледаним методама машинског учења те коришћењем библиотеке Scikit-Learn у решавању стварних проблема моделовања и предикције из домена климатских промена. | | | | |
| **Исход предмета**  Студенти су оспособљени да:   * користе библиотеке Numpy и Pandas у припреми и почетној обради скупова података, * користи библиотеку Matplotlib за визуалну интерпретацију података, * користи библиотеку Scikit-Learn за имплементацију надгледаних и ненадгледаних метода машинског учења, * примењује методе машинског учења на решавање проблема из домена климатских промена. | | | | |
| **Садржај предмета**  **Наука о подацима у Пајтону.** Понављање основна процедуралног и објектно-оријентисаног програмирања у Пајтону. Библиотека Numpy. Табеларна организација података. Библиотека Pandas. Дескриптивна статистика. Појам променљиве. Нумеричке и категоријалне променљиве. Очекивана вредност. Варијанса. Стандардна девијација. Медијана и мод. Распон и квартили.  **Визуализација података.** Билбиотека Matplotlib.  **Машинско учење.** Увод. Примери. Основни појмови. Надгледано и ненадгледано учење. Проблем регресије и класификације. Библиотека Scikit-Learn. Хиперпараметри и евалуација модела. Почетна обрада података. Одабир својстава. Наивни Бајесов класификатор. Линеарна регресија. Регуларизација. Логистичка регресија. Метод потпорних вектора. Случајне шуме. Редукција димензије. Кластеровање података.  **Примери из праксе**. Примене метода машинског учења. | | | | |
| **Литература**   1. Jake VanderPlas: “Python Data Science Handbook: Essential Tools for Working with Data“, O'Reilly Media, 1st edition (2017) 2. Aurélien Géron. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O’Reilly Media, 2nd edition, (2019) | | | | |
| **Број часова активне наставе** | **Теоријска настава:** 3 | | **Практична настава:** 0 | |
| **Методе извођења наставе**  Настава се изводи комбинацијом традиционалних и савремених метода: предавање наставника уз коришћење интерактивних презентација и рад на практичним задацима у рачунарским учионицама. | | | | |
| **Оцена знања (максимални број поена 100)** | | | | |
| **Предиспитне обавезе** | 70 поена | **Завршни испит** | | 30 поена |
| колоквијум-и | 50 поена | писмени испит | | 30 поена |
| семинарски рад | 20 поена |  | |  |