



УТВЪРДИЛ:

Декан

Дата

СОФИЙСКИ УНИВЕРСИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“

Факултет по славянски филологии

Специалност: (код и наименование)

--	--	--	--	--	--	--	--	--

.....
Магистърска програма: (код и наименование)

С	Л	Б	3	0				
---	---	---	---	---	--	--	--	--

Компютърна лингвистика. Интернет технологии в хуманитаристиката
(редовна форма на обучение)

УЧЕБНА ПРОГРАМА

Дисциплина:

--	--	--	--

Увод в машинното самообучение

Преподавател: проф. д-р Мария Нишева

Асистент:

Учебна заетост	Форма	Хорариум
Аудиторна заетост	Лекции	30 часа
	Семинарни упражнения	
	Практически упражнения	
Обща аудиторна заетост		
Извънаудиторна заетост	Реферат	
	Доклад/Презентация	
	Научно есе	
	Курсов учебен проект	30 часа
	Учебна екскурзия	
	Самостоятелна работа в библиотека или с ресурси	30 часа
Обща извънаудиторна заетост		
ОБЩА ЗАЕТОСТ		90 часа
Кредити аудиторна заетост		1
Кредити извънаудиторна заетост		2
ОБЩО ЕСТК		3

№	Формиране на оценката по дисциплината	% от оценката
1.	Курсов проект	50%
2.	Изпит	50%

Анотация на учебната дисциплина:

Курсът запознава студентите с основните принципи на функциониране на самообучаващите се интелигентни системи и с множество базови подходи в областта на машинното самообучение, сред които: класификационни дървета, покриващи правила, невронни мрежи, самообучение чрез запомняне, Бейсови методи за самообучение и др.

Изучават се и базови алгоритми, използвани за решаване на задачи от областта на извличането на закономерности от данни: подготовка и преобразуване на данни, клъстерен анализ, класификация, регресия и др.

Предварителни изисквания:

Изслушан курс „Основи на програмирането (Python)“.

Очаквани резултати:

Успешното завършване на курса ще даде на обучаемите:

- знания за основните принципи на функциониране на самообучаващите се интелигентни софтуерни системи;
- знания за базовите алгоритми, реализиращи основните подходи в областта на машинното самообучение;
- знания за основните типове задачи за извличане на закономерности от данни и начални умения за прилагане на базовите алгоритми за тяхното решаване.

Учебно съдържание

№	Тема:	Хорариум
1	Основни характеристики на интелигентните софтуерни системи. Интелигентни агенти. Типове интелигентни агенти. Предмет и цели на машинното самообучение (МС). Основни типове МС. Самообучаващи се интелигентни агенти	3 часа
2	Научаване на понятия – построяване на класификационно дърво (дърво на решенията). Избор на атрибути – печалба на информация	2 часа
3	Научаване на множество от правила чрез последователно покриване. Сравнение на разгледаните методи за научаване на множество от правила	3 часа
4	Изкуствени невронни мрежи (ИНМ). Двуслоен персептрон. Многослойни ИНМ. Алгоритъм за обучение	4 часа

	с обратно разпространение на грешката. Приложения при решаване на реални задачи	
5	Бейсов класификатор. Наивен Бейсов класификатор. Класификация на текстове	2 часа
6	Бейсови мрежи. Вероятностен извод. Научаване на Бейсови мрежи	2 часа
7	МС чрез запомняне. Алгоритъм на k най-близки съсед	2 часа
8	Основни понятия от областта на извличането на закономерности от данни (ИЗД). Модел на процеса на ИЗД. Типове задачи за ИЗД	2 часа
9	Разбиране на данни. Предварителна подготовка на данните. Намалване на обема на данните	3 часа
10	Клъстерен анализ. Основни методи за нейерархична и йерархична клъстеризация	3 часа
11	Предсказване чрез регресионни дървета	2 часа
12	Извличане на асоциативни правила	2 часа

Конспект за изпит

№	Въпрос
1	Предмет и цели на машинното самообучение (МС). Основни типове МС. Самообучаващи се интелигентни агенти
2	Построяване на класификационно дърво (дърво на решенията). Критерий за избор на атрибути
3	Научаване на множество от правила чрез последователно покриване
4	Изкуствени невронни мрежи (ИНМ). Двуслоен перцептрон
5	Многослойни ИНМ. Алгоритъм за обучение с обратно разпространение на грешката
6	Бейсов класификатор. Наивен Бейсов класификатор
7	Класификация на текстове
8	Бейсови мрежи. Вероятностен извод. Научаване на Бейсови мрежи
9	МС чрез запомняне. Алгоритъм на k най-близки съсед
10	Основни понятия в извличането на закономерности от данни (ИЗД). Модел на процеса на ИЗД. Типове задачи за ИЗД
11	Разбиране на данни. Предварителна обработка на данните
12	Клъстеризация. Основни типове методи за клъстеризация. Метод k -means
13	Предсказване чрез регресионни дървета
14	Извличане на асоциативни правила

Библиография

Основна:

1. Mitchell, T. Machine Learning. McGraw-Hill, 1997.
2. Witten, I., E. Frank, M. Hall, C. Pal. Data Mining: Practical Machine Learning Tools and Techniques (4th ed.). Morgan Kaufmann, 2017.

Допълнителна:

1. Azzalini, A., V. Scarpa. Data Analysis and Data Mining: An Introduction. Oxford University Press, 2012.
2. Агре, Г., З. Марков, Д. Дочев. Увод в машинното самообучение. СОФТЕХ, 2001.
3. Кирова, Т. Невронни мрежи: Основни архитектури и обучаващи алгоритми. СОФТЕХ, София, 1995.

Дата: 16.02.2022 г.

Съставил:

проф. д-р Мария Нишева