

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«Київський політехнічний інститут імені Ігоря Сікорського»

Затверджую



Протокол Виконавчої комісії

Михайло
ЗГУРОВСЬКИЙ

28.04.2023

дата

Навчально-науковий інститут прикладного системного аналізу
повна назва факультету/навчально-наукового інституту

ПРОГРАМА
комплексного фахового випробування
для вступу на освітньо-професійну програму підготовки магістра
«Системи і методи штучного інтелекту»
за спеціальністю 122 Комп'ютерні науки

Програму ухвалено:

Вченою Радою Навчально-наукового інституту
прикладного системного аналізу

Протокол № 3 від «27» березня 2023 р.

Заступник Голови Вченої Ради

Віктор РОМАНЕНКО

ВСТУП

Програма комплексного фахового випробування для вступу на освітню програму підготовки магістра «Системи і методи штучного інтелекту» за спеціальністю 122 «Комп'ютерні науки» (далі — Програма) призначена для отримання досвіду самостійної роботи абітурієнта з підготовки до екзамену.

Метою програми є формування у абітурієнтів здатностей:

ознайомитися із предметними питаннями курсів навчальних дисциплін, що включені в екзаменаційні білети;

опрацювати підручники, навчальні посібники та інші інформаційно-літературні джерела предметної області знання;

осмислити, упорядити і систематизувати засвоєні теоретичні знання і практичні навички;

вмотивовано виконати роботу на екзамені, продемонструвавши певний рівень засвоєння структури та змісту навчальних дисциплін в результаті підготовки.

Перелік цієї Програми складають навчальні дисципліни базової підготовки та, відповідно, належать до циклу загальної підготовки навчального плану підготовки бакалавра за освітньо-професійною програмою «Системи і методи штучного інтелекту» спеціальності 122 «Комп'ютерні науки»:

1) Методи та системи штучного інтелекту

2) Вступ до інтелектуального аналізу даних

Вступний екзамен проводиться чотири академічних години без перерви (180 хвилин), в аудиторному приміщенні випускової кафедри штучного інтелекту КПІ ім. Ігоря Сікорського, за методом «одержання екзаменаційного білету— повернення письмової роботи».

Завданням на екзамені є розв'язання задач екзаменаційного білету. Екзаменаційний білет містить три теоретичних і одне тестове завдання. Диференціації робочого часу, відведеного на виконання кожного завдання, немає. Фіксується час початку і закінчення роботи.

ПЕРЕЛІК НАВЧАЛЬНОГО МАТЕРІАЛУ, ЩО ВІНОСИТЬСЯ НА КОМПЛЕКСНЕ ФАХОВЕ ВИПРОБУВАННЯ

1. Методи та системи штучного інтелекту

НЕЙРОННІ МЕРЕЖІ ТА ЇХ ОСОБЛИВОСТІ

Переваги нейронних мереж. Математичні моделі нейронів. Типи функцій активації. Стохастична модель нейрона. Представлення нейронних мереж за допомогою направлених графів. Зворотній зв'язок. Архітектура нейронних мереж. Представлення знань. Штучний інтелект та нейронні мережі.

ПРОЦЕСИ НАВЧАННЯ

Класифікація методів навчання. Навчання, що базується на корекції помилок. Навчання з урахуванням пам'яті. Навчання Хебба. Конку rentне навчання. Навчання Больцмана. Завдання привласнення коефіцієнтів довіри. Навчання з учителем. Навчання без учителя. Навчання із підкріпленням. Статистична природа процесу навчання. Принцип мінімізації емпіричного ризику. Мінімізація структурного ризику.

ОДНОШАРОВИЙ ПЕРСЕПТРОН

Методи безумовної оптимізації. Метод якнайшвидшого спуску. Метод Ньютона. Метод Гауса-Ньютона. Алгоритм мінімізації середньоквадратичної помилки. Теорема про збіжність персептрону.

БАГАТОШАРОВИЙ ПЕРСЕПТРОН

Три відмінні ознаки персептрону. Обчислювальна ефективність методу зворотного поширення, Архітектура багатошарового персептрона. Два типи сигналів багатошарового персептрону. Алгоритм зворотного розповсюдження, обчислення локального градієнта. Дельта-правило. Алгоритм зворотного розповсюдження вихідного вузла. Алгоритм зворотного поширення прихованого вузла. Завдання призначення коефіцієнтів довіри. Два проходи обчислень в алгоритмі зворотного розповсюдження. Швидкість навчання алгоритму зворотного розповсюдження. Послідовний режим (sequential mode) навчання за методом зворотного розповсюдження. Пакетний режим (batch mode) навчання за методом зворотного поширення. Критерій зупинки. Завдання XOR. Евристичні рекомендації щодо поліпшення роботи алгоритму зворотного поширення. Представлення виходу та вирішальне правило. Байєсовський критерій оптимальної класифікації. Оптимальна кількість прихованих нейронів. Оптимальне навчання та константа моменту. Оцінка оптимальної архітектури мережі. Вилучення ознак. Зв'язок із лінійним дискримінантом Фішера. Граф чутливості (sensitivity graph). Матриця якобіана, гесіан. Узагальнення, перенавчання (overtraining чи overfitting). Достатній обсяг прикладів навчання для коректного узагальнення, VC вимір. Теорема про універсальну апроксимацію. "Прокляття розмірності". Локальні та глобальні ознаки. Вибір моделі на основі перехресної перевірки. Метод навчання багатошарового персептрону з раннім зупиненням. Варіанти реалізації методу перехресної перевірки. Регуляризація складності при структурно-параметричному синтезі персептрона (зниження та вилучення ваг). Спрощення структури мережі на основі гесіану. Процедура оптимального ступеня пошкодження (optima brain damage OBD). Процедурі оптимальної хірургії мозку (optimal brain surgeon OBS). Переваги та обмеження навчання методом зворотно поширення. Властивість складності багатошарового персептрона. Вилучення ознак багатошаровим персептроном. Прискорення збіжності процесу навчання методом зворотного

поширення. Навчання з учителем як завдання оптимізації. Метод сполучених градієнтів.

Література [1-7]

2 Вступ до інтелектуального аналізу даних

Різновиди задач Data Mining

Методи препроцесингу інформації. Нормалізація, масштабування, збільшення інформативності вхідних факторів.

Методи препроцесингу інформації. Алгоритм вибілювання входів

Методи препроцесингу інформації. Виявлення мультиколінеарності факторів

Методи препроцесингу інформації: тестування і усунення гетероскедастичності та автокореляції

Метод головних компонентів

Аналітико-евристичні алгоритми визначення вагомих інформативних ознак: алгоритми Add, Del та їх поєднання

Аналітико-евристичні алгоритми визначення вагомих інформативних ознак: алгоритм випадкового пошуку з адаптацією, методика Vox-counting

Методи кластеризації: формальна постановка задачі, методи визначення оптимальної кількості кластерів

Методи кластеризації: методи агломеративної кластеризації

Методи кластеризації: ітеративні методи кластеризації

Методи кластеризації: методи засновані на гіпотезі щільності

Методи кластеризації: графові методи кластеризації

Методи кластеризації: карти Кохонена

Методи кластеризації: EM-алгоритм

Методи пошуку асоціацій. Характеристики асоціативних правил

Методи пошуку асоціацій. Алгоритм Apriori

Відновлення пропусків у даних. Формальна постановка задачі. Найпростіші методи відновлення пропусків

Відновлення пропусків у даних. Методи Resampling

Відновлення пропусків у даних. Метод Бартлетта

Відновлення пропусків у даних. Алгоритм Zet

Еволюційний метод відновлення пропусків

Задачі класифікації: формальна постановка. Наївний Байес

Задачі класифікації: метод опорних векторів

Задачі класифікації: бінарні дерева рішень
Ансамблеві деревовидні моделі: випадковий ліс і алгоритм градієнтного бустингу.

Література [1,2, 8]

ПРИКІНЦЕВІ ПОЛОЖЕННЯ

Користування допоміжним матеріалом на екзамені

— Забороняється

Критерії оцінювання (за системою ECTS, столбальна шкала)

Письмова відповідь на запитання або розв'язання кожної задачі оцінюється за такими критеріями:

95—100	—	Відповідь на теоретичне питання дана правильно, в повному обсязі, задачу розв'язано повністю, вірно
85—94	—	Відповідь на теоретичне питання дана правильно, але в неповному обсязі, задачу розв'язано вірно, відповідь правильна, але наявними є один-два недоліки (наявними є деякі методичні помилки, порушено послідовність викладок тощо)
75—84	—	Відповідь на теоретичне питання частково відповідає змісту питання але проявлено розуміння проблеми та причинно-наслідковий зв'язок в розв'язанні проблеми, задачу розв'язано вірно, але відповідь неправильна
65—74	—	Проявлено розуміння загальних понять, причинно-наслідковий зв'язок в розв'язанні проблеми не відчувається, задачу розв'язано неповністю, але намічено правильний хід розв'язування
60—64	—	Відповідь на теоретичне питання не відповідає суті питання, а показує тільки загальні знання за дисципліною, задачу не розв'язано, але наведено формули або твердження, що можуть бути використані при розв'язуванні задачі
менше 60	—	Відповідь на теоретичне питання не надано взагалі, задачу не розв'язано

Результат роботи обчислюється як середнє арифметичне оцінок, що їх отримано за кожну відповідь на теоретичне питання або розв'язання задачі і заокруглюється до цілих.

Відповідно до «Правил прийому до КПІ ім. Ігоря Сікорського в 2020 році» є необхідним перерахунок оцінки рейтингової системи оцінювання (60...100 балів) в шкалу ЄВІ (100...200 балів).

Таблиця. Відповідність оцінок рейтинговій системі оцінювання (PCO)

Таблиця відповідності оцінок PCO (60...100 балів)
оцінкам 200-бальної шкали (100...200 балів)

шкала PCO	шкала 100...200	шкала PCO	шкала 100...200	шкала PCO	шкала 100...200	шкала PCO	шкала 100...200
60	100	70	140	80	160	90	180
61	105	71	142	81	162	91	182
62	110	72	144	82	164	92	184
63	115	73	146	83	166	93	186
64	120	74	148	84	168	94	188
65	125	75	150	85	170	95	190
66	128	76	152	86	172	96	192
67	131	77	154	87	174	97	194
68	134	78	156	88	176	98	196
69	137	79	158	89	178	99	198
						100	200

Приклад типового завдання комплексного фахового випробування

Білет № 0

1. В чому полягає Обчислювальна ефективність методу зворотного поширення?
2. Обґрунтуйте необхідність застосування регуляризації складності при структурно-параметричному синтезі персептрона (зниження та вилучення ваг).
3. Розкрийте суттєвий зміст методів кластеризації, які засновані на гіпотезі щільності. Дайте їх порівняльну оцінку.
4. Встановіть відповідність:

<p>Яка з формул приводить значення даних до:</p> <ol style="list-style-type: none"> 1. відрізка [0,1] 2. відрізка [-1,1] 3. стандартно розподілених 4. більш рівномірно розподілених значень 	<p>А) $x'_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}$</p> <p>Б) $x'_i = \frac{1}{1 + e^{-x_i}}$</p> <p>В) $x'_i = \frac{2(x_i - \min(x_i))}{\max(x_i) - \min(x_i)} - 1$</p> <p>Г) $x'_i = \frac{\max(x_i) - x_i}{\max(x_i) - \min(x_i)}$</p>
--	---

	$x'_i = \frac{x_i - \bar{x}_i}{\sigma_{x_i}}$ Д)
--	--

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

- [1] Michael Z. Zgurovsky Artificial Intelligence Systems Based on Hybrid Neural Networks / Michael Z. Zgurovsky, Victor M. Sineglazov, Olena I. Chumachenko // Springer <https://link.springer.com/book/10.1007/978-3-030-48453-8>. Customer can order it ia <https://www.springer.com/gp/book/9783030484521>. – 510 p.
- [2] Simon Haykin Neural Networks A Comprehensive Foundation – Prentice Hall Upper Saddle River, New Jersey 07458 – 1104 p.
- [3] Синєглазов В.М. Інтелектуальне управління дорожнім рухом / В.М. Синєглазов, О.І. Чумаченко, В.С. Горбатюк. К.: Освіта України, 2013. – 194 с.
- [4] Синєглазов В.М. Інтелектуальні методи прогнозування / В.М. Синєглазов, О.І. Чумаченко, В.С. Горбатюк. К.: Освіта України, 2013. – 236 с.
- [5] Булгакова, О.С. Методи та системи штучного інтелекту: теорія та практика / О.С. Булгакова, В.В. Зосімов, В.О. Поздєєв: навч. посіб. – Херсон: Олді-Плюс, 2020. – 353 с.
- [6] Засоби штучного інтелекту / Р.О. Ткаченко ... [та ін.]: навч. посіб. – Львів: вид-во Львівської політехніки, 2014. – 203 с.
- [7] Нікольський, Юрій Володимирович. Системи штучного інтелекту / Ю.В. Нікольський, В.В. Пасічник, Ю.М. Щербина: навч. посіб. – Львів: вид-во "Магнолія-2006", 2021. – 278 с.
- [8] Купенко О.В. Інтелектуальний аналіз даних. Конспект лекцій. – Дніпро: вид-во Національного гірничого університету, 2019. – 86 с.

Розробники програми:

Чумаченко О.І., завідувачка кафедри штучного інтелекту, докт. техн. наук, проф.

Олена ЧУМАЧЕНКО

Синєглазов В.М., професор кафедри штучного інтелекту, докт. техн. наук, проф.

Віктор СИНЄГЛАЗОВ

Купенко О.В., професор кафедри штучного інтелекту, докт. техн. наук, доцент

Ольга КУПЕНКО