

НАЦІОНАЛЬНИЙ ТРАНСПОРТНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ТРАНСПОРТНИХ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ

ЗАТВЕРДЖУЮ:

Завідувач кафедри інформаційних систем і технологій

проф. В.В. Гавриленко _____
_____ 2020 р.

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
ОКП 27 ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ

Рівень вищої освіти	Перший (бакалаврський)
Спеціальність	121 Інженерія програмного забезпечення
Освітня програма	Інженерія програмного забезпечення http://vstup.ntu.edu.ua//osvitprog/FTIT/121IPZ_2020.pdf
Тип дисципліни	Обов'язкова
Форма навчання	Денна
Семестр	7-й семестр навчального плану
Розробник	Зубрецька Н. А. д.т.н., проф., Федін С.С., д.т.н., проф.
Викладач	Зубрецька Наталія Анатоліївна д.т.н., проф. е-mail викладача: zubr2767@gmail.com
Доступ до матеріалів	https://classroom.google.com/c/MTU4NDIzOTgzOTcw https://classroom.google.com/c/MTU4NDIzOTgzOTg0Njc1
Кафедра	інформаційних систем і технологій Тел. кафедри: +38 (044) 280-70-66 Веб-сайт кафедри: http://kist.ntu.edu.ua/
Гарант освітньої програми	к.ф.-м. н., доцент Вітер Михайло Богданович

1. АНОТАЦІЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Мета вивчення дисципліни – набуття ключових фахових компетентностей, теоретичних знань і практичних навичок з інтелектуального аналізу даних у різних сферах професійної діяльності.

Предметом вивчення дисципліни є технології, методи та засоби інтелектуального аналізу даних.

Завдання вивчення дисципліни:

- оволодіння основними поняттями інтелектуального аналізу даних;
- ознайомлення з новітніми інформаційними технологіями аналітичної обробки інформації;
- набуття практичних навичок використання методів і засобів інтелектуального аналізу даних.

Мова викладання: українська.

2. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Навчальна дисципліна покликана допомогти студенту отримати:

знання основних понять, методів, засобів, моделей та алгоритмів ІАД;

розуміння принципів застосування технологій інтелектуального аналізу даних;

уміння вільно орієнтуватися на сучасному ринку аналітичних програмних продуктів; практично застосовувати методи консолідації, трансформації, візуалізації, оцінки якості, очищення та предобробки даних для якісної підготовки даних до аналізу; використовувати ПЗ для використання процедур інтелектуального аналізу при обробці інформації;

здатність використовувати можливості вітчизняних і зарубіжних універсальних програмних засобів і аналітичних платформ для пошуку закономірностей, зв'язків, правил, знань в електронних масивах даних.

КОМПЕТЕНТНОСТІ

Загальні компетентності

K01. Здатність до абстрактного мислення, аналізу та синтезу.

K02. Здатність застосовувати знання у практичних ситуаціях.

K05. Здатність вчитися і оволодівати сучасними знаннями.

K06. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

Спеціальні (фахові, предметні) компетентності

K19. Володіння знаннями про інформаційні моделі даних, здатність створювати програмне забезпечення для зберігання, видобування та опрацювання даних.

K20. Здатність застосовувати фундаментальні і міждисциплінарні знання для успішного розв'язання завдань інженерії програмного забезпечення.

K21. Здатність оцінювати і враховувати економічні, соціальні, технологічні та екологічні чинники, що впливають на сферу професійної діяльності.

K26. Здатність до алгоритмічного та логічного мислення

K27. Здатність застосовувати інформаційні технології для моделювання транспортних процесів.

K31. Здатність застосовувати на практиці сучасні інформаційні технології відповідно до розв'язуваних прикладних завдань.

K33. Здатність застосовувати методи та алгоритми теорії розпізнавання образів для розв'язання прикладних завдань.

ПРОГРАМНІ РЕЗУЛЬТАТИ НАВЧАННЯ

ПР01. Аналізувати, цілеспрямовано шукати і вибирати необхідні для вирішення професійних завдань інформаційно-довідникові ресурси і знання з урахуванням сучасних досягнень науки і техніки.

ПР07. Знати і застосовувати на практиці фундаментальні концепції, парадигми і основні принципи функціонування мовних, інструментальних і обчислювальних засобів інженерії програмного забезпечення.

ПР09. Знати та вміти використовувати методи та засоби збору, формулювання та аналізу вимог до програмного забезпечення.

ПР18. Знати та вміти застосовувати інформаційні технології обробки, зберігання та передачі даних.

ПР25. Аналізувати, оцінювати і вибирати інструментальні засоби, технології, алгоритмічні та програмні рішення для моделювання транспортних процесів.

ПР29. Вміти вибирати та застосовувати на практиці сучасні інформаційні технології відповідно до розв'язуваних прикладних завдань.

ПР30. Вміти застосовувати методи та алгоритми теорії розпізнавання образів для розв'язання прикладних завдань.

СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Види робіт за навчальним планом	Години
Аудиторні заняття, у т.ч.:	48
Лекції	16
Лабораторні роботи	32
Практичні заняття	–
Самостійна робота, у т.ч.:	72
Підготовка до аудиторних занять	24
Підготовка до контрольних заходів	4
Виконання курсової роботи	30
Опрацювання питань програми, які не викладаються на лекціях	10
Підготовка до екзамену	4
Всього:	120 (4 кредити)
Форма підсумкового контролю	Екзамен

ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, ВИКОРИСТАННЯ ЯКИХ ПЕРЕДБАЧАЄ НАВЧАЛЬНА ДИСЦИПЛІНА

№	Найменування
1.	Опорний конспект лекцій
2.	Навчальні посібники
3.	Силабус
4.	Програмне забезпечення: STATISTICA
5.	Комплект контрольних завдань для поточного оцінювання навчальних досягнень
6.	Засоби підсумкового контролю (комплект завдань для підсумкового контролю)

3. ОРГАНІЗАЦІЯ НАВЧАННЯ

Назви модулів і тем лекцій	Кількість годин			
	усього	у тому числі		
		лекції	Лабор.	самостійна робота
модуль 1. Технології аналізу та трансформації даних методами Data Mining				
1. Технології аналізу даних	10	2	8	
2. Трансформація даних	10	2		8
модуль 2. Попередня обробка, візуалізація та оцінювання якості даних				
3. Візуалізація даних	10	2	8	
4. Статистичні методи аналізу даних (Оцінка якості та попередня обробка даних)	10	2		8
модуль 3. Поглиблений аналіз даних				
5. Пошук асоціативних правил	10	2	8	
3. Методи кластеризації	10	2		8
модуль 4. Прогнозування та машинне навчання стемі				
4. Аналіз та прогнозування часових рядів	10	2	8	
8. Класифікація, регресія, машинне навчання	20	2		18
Курсова робота	30			30
Усього годин за рік	120	16	32	72

ЛАБОРАТОРНІ РОБОТИ

№	Назва теми	Кількість годин
1	Розвідувальний аналіз статистичних даних	8
2	Порівняльний аналіз статистичних даних	8
3	Кореляційно-регресійний аналіз даних у системі STATISTICA	8
4	Створення інформаційно-аналітичних моделей для аналізу даних і прогнозування	8
Всього		32

КУРСОВА РОБОТА

№	Етапи виконання роботи	Термін (тижні)	Кількість годин
1.	Вибір напрямку дослідження	1	2
2.	Пошук і опрацювання літератури	2	4
3.	Формулювання та затвердження теми	1	2
4.	Складання плану курсової роботи	2	4
5.	Поглиблений аналіз об'єкту дослідження	4	6
6.	Формування тексту роботи і списку літератури	4	6
7.	Оформлення пояснювальної записки та захист	2	6
Всього		16	30

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

- Боровиков В.П. Популярное введение в современный анализ данных в системе STATISTICA. Учеб.пособие для вузов. – М.: Горячая линия-Телеком, 2013. – 288 с
- Кацко И.А., Паклин Н.Б. Практикум по анализу данных на компьютере / Под ред. Г.В. Гореловой. – М.: КолосС, 2009. – 278 с.
- Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных: OLAP и Data Mining. – СПб.:БХВ –Петербург, 2004. – 336 с.
- Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных: OLAP и Data Mining. – СПб.:БХВ –Петербург, 2004. – 336 с.
- Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Технологии анализа данных. Data Mining, Visual Mining, Text Mining, OLAP . – СПб.: БХВ –Петербург, 2007. – 384 с.
- Петрунин Ю.Ю. Информационные технологии анализа данных. Data Analysis – 2-е издание. – М.: КДУ, 2010. – 292 с.
- Нейронные сети. STATISTICA Neural Networks: Методология и технологии современного анализа данных / Под редакцией В.П. Боровикова. – 2-е изд., перераб. и доп. – М.: Горячая линия – Телеком, 2008. – 392 с.
- Корнеев В.В., Гареев А.Ф., Васютин С.В., Райх В.В. Базы данных. Интеллектуальная обработка информации. – М.: Нолидж, 2000. – 352 с.
- Осовский С. Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского.- М.: Финансы и статистика, 2002. - 344 с.
- Рутковская Д., Пилинский м., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д. Рудинского. – М.: Горячая линия – Телеком, 2006. – 452 с.

Електронні ресурси

www.basegroup.ru
www.statsoft.ru
<http://www.sas.com/>
<http://www.predictivesolutions.ru/>
<http://www.kdnuggets.com/>
<http://statlab.kubsu.ru/node/1>
<http://itt-pgups.ru/>

Доступ до матеріалів щодо структури дисципліни в робочому навчальному плані та освітній програмі, анотації лекцій, лабораторних робіт, тематики та методичних вказівок до виконання курсових робіт, самотійної роботи, питань до модульних контрольних робіт, підсумкового контролю та ін. можна отримати за посиланнями:

<http://vstup.ntu.edu.ua/navch-plan/FTIT-PR.pdf>

<http://vstup.ntu.edu.ua/osvitprog/FTIT/121IPZ.pdf>

для групи ПР-4-1 <https://classroom.google.com/c/MTU4NDIzOTgzOTcw>

для групи ПР-4-2 <https://classroom.google.com/c/MTU4NDIzOTg0Njc1>

4. ПОРЯДОК ТА КРИТЕРІЇ ОЦІНЮВАННЯ

Розподіл балів, які отримують студенти за модулі

Поточне оцінювання модулів							Екзамен	Сума
	M1	M2	MK1	M3	M4	MK2		
	10	10	10	10	10	10	40	100
Присутність на лекціях	2	2		2	2			
Присутність на ЛР	2	2		2	2			
Виконання та захист ЛР	6	6		6	6			

Модульна оцінка (максимальна кількість балів – 30) складається із:

- присутності студента на лекціях (максимальна кількість балів – 4);
- присутності на лабораторних заняттях (максимальна кількість балів – 4);
- виконання та захисту лабораторних робіт (максимальна кількість балів – 12);
- модульної контрольної роботи (максимальна кількість балів – 10).

Модульна контрольна робота складається з 15 питань теоретичного курсу.

Максимальна кількість балів за кожне питання:

- за повністю розкритою відповіддю на питання студент одержує 2 бали;
- якщо у відповіді не повністю розкрито сутність питання та допущені невірні тлумачення, студент одержує 1 бал;
- якщо студент не надав відповідь на питання, повністю не виконано завдання, або допущено принципові помилки, – студент одержує 0 балів.

Екзамен (максимальна оцінка за екзамен – 40 балів). Екзаменаційне завдання складається з трьох питань теоретичного курсу та 1 практичного завдання.

Максимальна кількість балів за кожне питання та завдання:

- за повністю розкритою відповіддю на питання студент одержує 10 балів;
- якщо студент дав відповідь на питання і виконав завдання, допустивши не принципові помилки, студент одержує 7 балів;

- якщо у відповіді не повністю розкрито сутність питання та допущені невірні тлумачення, студент одержує 3 бали;
- якщо студент не надав відповідь на питання, не виконав завдання, або виконав завдання з принциповими помилками, – одержує 0 балів.

Підсумкова оцінка з дисципліни визначається як сума балів за всі види діяльності.

5. ПОЛІТИКА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Форми організації освітнього процесу, види навчальних занять і оцінювання результатів навчання регламентуються Положенням про організацію освітнього процесу в Національному транспортному університеті та Положенням про систему внутрішнього забезпечення якості вищої освіти.

Інформація про мету, завдання, структуру і порядок вивчення навчальної дисципліни надається здобувачам на початку семестру у вигляді **навчально-методичного комплексу (НМК)**, склад якого регламентується Переліком навчально-методичного забезпечення дисциплін.

Політика виставлення оцінок: кожна оцінка виставляється відповідно до розроблених викладачем та заздалегідь оголошених студентам критеріїв, а також мотивується в індивідуальному порядку на вимогу студента; у випадку не виконання студентом усіх передбачених навчальним планом видів занять (лабораторних робіт, курсової роботи) до екзамену він не допускається; пропущені заняття обов'язково мають бути відпрацьовані.

Відвідування є обов'язковим (за винятком випадків, коли існує поважна причина, наприклад, хвороба чи дозвіл працівників деканату). У деяких випадках можливе зарахування окремих тем, модулів дисципліни, що регламентується Тимчасовим положенням про порядок визнання результатів навчання, набутих студентами Національного транспортного університету у неформальній/інформальній освіті.

Назви модулів і тем лекцій	Рекомендовані ресурси
модуль 1. Технології аналізу та трансформації даних методами Data Mining	
1. Технології аналізу даних	https://courses.prometheus.org.ua/courses/IRF/Stat101/2016_T3/about
2. Трансформація даних	
модуль 2. Попередня обробка, візуалізація та оцінювання якості даних	
3. Візуалізація даних	https://courses.prometheus.org.ua/courses/IRF/DV101/2016_T3/about
4. Статистичні методи аналізу даних	https://courses.prometheus.org.ua/courses/IRF/Stat101/2016_T3/about
модуль 3. Поглиблений аналіз даних	
5. Пошук асоціативних правил	Sequence models. Deeplearning.ai/Coursera Convolutional Neural Networks. Deeplearning.ai/Coursera
3. Методи кластеризації	Browser-based Models with TensorFlow.js. Deeplearning.ai/Coursera
модуль 4. Прогнозування та машинне навчання стемі	
8. Класифікація, регресія, машинне навчання	Машинне навчання https://courses.prometheus.org.ua/courses/IRF/ML101/2016_T3/about Machine Learning. Стенфордський університет/ Coursera Введение в Data Science и машинное обучение/Bioinformatics Institute/Stepik.ru

Порядок зарахування пропущених занять. Відпрацювання пропущеного заняття з лекційного курсу здійснюється шляхом підготовки і захисту реферату за відповідною темою у вигляді презентації відповідно до графіку консультацій викладача. Відпрацювання пропущеного лабораторного заняття здійснюється шляхом самостійного виконання завдання і його захисту відповідно до графіку консультацій викладача.

Політика академічної доброчесності. Плагіат та інші форми нечесної роботи неприпустимі. Всі індивідуальні завдання та курсову роботу студент має виконати самостійно із використанням рекомендованої літератури й отриманих знань та навичок. Цитування в письмових роботах допускається тільки із відповідним посиланням на авторський текст. Недопустимі підказки і списування у ході захисту лабораторних робіт, на контрольних роботах, на іспиті. Дотримання академічної доброчесності студентів і викладачів регламентується [Положенням про систему забезпечення академічної доброчесності педагогічними, науково-педагогічними, науковими працівниками та здобувачами вищої освіти в Національному транспортному університеті](#) та вимогами [Антикорупційної програми](#).

Норми академічної етики – дисциплінованість; дотримання субординації; чесність; відповідальність; робота в аудиторії з відключеними мобільними телефонами – задекларовані у [Кодексі етики академічних взаємовідносин та доброчесності Національного транспортного університету](#).

При виконанні лабораторних робіт студент може користуватися ноутбуками. Проте під час лекційних занять та обговорення завдань лабораторних робіт не слід використовувати ноутбуки, смартфони, планшети чи комп'ютери. Це відволікає викладача і студентів групи та перешкоджає навчальному процесу. Якщо ви використовуєте свій ноутбук чи телефон для аудіо- чи відеозапису, необхідно заздалегідь отримати дозвіл викладача. Повага один до одного дає можливість ефективніше досягати поставлених командних результатів.

Конфліктні ситуації мають відкрито обговорюватись в академічних групах з викладачем, необхідно бути взаємно толерантним, поважати думку іншого. Для запобігання конфліктних ситуацій в НТУ є можливість скористатися «Скринькою довіри» відповідно до [Положення про функціонування у Національному транспортному університеті «Скриньки довіри» з питань запобігання виникненню конфліктних ситуацій](#),