

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО»
Фізико-технічний інститут
Кафедра математичного моделювання та аналізу даних

МОДЕЛІ ТА МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ГЕТЕРОГЕННИХ ДАНИХ

СЕРТИФІКАТНА ПРОГРАМА

для другого (магістерського) рівня вищої освіти
за освітньою програмою «Математичні методи
моделювання, розпізнавання образів та комп'ютерного
зору»
спеціальності 113 Прикладна математика

*Ухвалено Методичною радою
КПІ ім. Ігоря Сікорського
від . .2021 р., протокол № _____*

*Введено в дію наказом
від . .2021 р., № / /2021*

Розробники сертифікатної програми:

Куссуль Наталя Миколаївна, доктор технічних наук, професор, в.о. завідувача математичного моделювання та аналізу даних

Шелестов Андрій Юрійович, доктор технічних наук, професор, професор кафедри математичного моделювання та аналізу даних

ОПИС СЕРТИФІКАТНОЇ ПРОГРАМИ

1. Загальна інформація

Назва сертифікатної програми	Моделі та методи інтелектуального аналізу гетерогенних даних
Рівень вищої освіти	Другий (магістерський)
Галузь знань	11 Математика та статистика
Спеціальність	113 Прикладна математика
Освітня програма	Математичні методи моделювання, розпізнавання образів та комп'ютерного зору
Факультет / Інститут	Фізико-технічний інститут
Кафедра	Математичного моделювання та аналізу даних
Обсяг сертифікатної програми	14 кредитів ЄКТС
Мова викладання	Українська
Документ про опанування сертифікатної програми	Сертифікат встановленого зразка КПІ ім. Ігоря Сікорського
Термін дії сертифікатної програми	Безстроково
Інтернет- адреса постійного розміщення сертифікатної програми	http://mmda.ipt.kpi.ua/

2. Мета сертифікатної програми

Сертифікатну програму «Моделі та методи інтелектуального аналізу гетерогенних даних» розроблено як профілізаційну складову освітньої програми за спеціальністю 113 «Прикладна математика» галузі знань 11 «Математика та статистика» другого (магістерського) рівня здобувачів вищої освіти в Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського».

Сертифікатну програму призначено для задоволення освітніх потреб студентів університету та зовнішніх слухачів, здатних розв'язувати складні спеціалізовані задачі та практичні проблеми із застосування математичних моделей та інтелектуальних методів обробки гетерогенних даних великого об'єму у хмарних середовищах.

Мета сертифікатної програми полягає в підготовці фахівців на основі поглиблення фахових компетентностей освітньої програми за спеціальністю 113 «Прикладна математика» у напрямках:

- розроблення математичних методів та моделей машинного та глибокого навчання для розв'язання задач на різнорідних даних;
- застосування методів оптимізації, розпізнавання образів, статистики та глибокого навчання для інтелектуального аналізу великих об'ємів різнорідних даних;
- розв'язання задач прогнозування, аналізу ризиків на основі гетерогенних даних різної природи;
- застосування інтелектуального аналізу гетерогенних даних для задач бізнес-аналізу та цифровізації економіки;
- розроблення інформаційних технологій для аналізу гетерогенних даних великого об'єму у хмарному середовищі.

3. Особливості участі слухачів Сертифікатної програми

3.1. Сертифікатна програма (СП) «Моделі та методи інтелектуального аналізу гетерогенних даних» представляє спеціалізований комплекс пов'язаних між собою дисциплін (освітніх компонентів) встановленої тривалості, який передбачає професійно спрямовану підготовку слухачів (студентів кафедри математичного моделювання та аналізу даних (ММАД) Фізико-технічного інституту та інших факультетів Національного технічного університету України КПІ ім. Ігоря Сікорського, а також зовнішніх слухачів).

3.2. Освітні компоненти сертифікатної програми складаються з вибіркового дисциплін другого (магістерського) рівня вищої освіти за спеціальністю «Прикладна математика» загальним обсягом 14 кредитів.

3.3. Запис слухачів на сертифікатну програму здійснюється на основі поданої заяви у встановлені кафедрою терміни. Запис здобувачів вищої освіти КПІ ім. Ігоря Сікорського на дисципліни СП проводиться у встановлені терміни запису на вибіркові дисципліни. Запис зовнішніх слухачів на дисципліни СП забезпечується кафедрою ММАД і здійснюється на весь обсяг СП через подання зовнішніми слухачами відповідної заяви, на підставі якої слухач зараховується до групи з проходження СП.

3.4. Зарахування слухачів на СП здійснюється за розпорядженням заступника директора з науково-педагогічної роботи Фізико-технічного інституту Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

3.5. Сертифікатна програма «Моделі та методи інтелектуального аналізу гетерогенних даних» для студентів КПІ ім. Ігоря Сікорського може бути реалізована в межах освітньої програми «Математичні методи моделювання, розпізнавання образів та комп'ютерного зору», за якою він навчається шляхом формування індивідуальної освітньої траєкторії з вибором всіх дисциплін, які пропонуються в межах даної СП.

3.6. Забезпечення цільової аудиторії передбачає формування групи з числа студентів кафедри ММАД, а також зовнішніх слухачів.

3.7. СП «Моделі та методи інтелектуального аналізу гетерогенних даних» надається студентам кафедри ММАД, іншим здобувачам вищої освіти університету та на платній основі для зовнішніх слухачів.

3.8. За результатами опанування слухачем освітніх компонентів сертифікатної програми, набуття певних професійних знань, умінь і навичок на основі здачі відповідних екзаменів та заліків видається сертифікат КПІ ім. Ігоря Сікорського за підписом проректора. Інформація про опанування сертифікатної програми може зазначатись в додатках до диплому європейського зразка.

4. Компетентності та очікувані результати навчання

Сертифікатну програму запроваджено як профілізаційну складову освітньої програми «Математичні методи моделювання, розпізнавання образів та комп'ютерного зору», для задоволення освітніх потреб здобувачів – формування ними індивідуальної траєкторії здобуття вищої освіти.

Сертифікатна програма передбачає підвищення рівня сформованості спеціальних (фахових) компетентностей за спеціальністю, посилення професійної підготовки за освітньою програмою.

Сертифікатна програма спрямована на засвоєння слухачами особливостей використання сучасних моделей та методів інтелектуального аналізу гетерогенних даних. Вона наповнена унікальним контентом та авторськими курсами, які характеризуються практичністю та актуальністю інформації, що дозволяє отримати додаткові знання та навички, розширити коло

кар'єрних можливостей в сфері використання моделей та методів інтелектуального аналізу гетерогенних даних з різних джерел та їх спільного використання для розв'язання важливих прикладних задач.

Фахові компетентності сертифікатної програми

ФК 1	Здатність аналізувати якість великих об'ємів різнорідних даних, здійснювати попередню обробку та злиття даних різної природи, зокрема результатів моделювання, для задач інтелектуального аналізу.
ФК 2	Здатність розробляти математичні методи та технології машинного та глибокого навчання, управління знаннями та інтелектуального аналізу даних для розв'язання задач цифровізації із різних предметних областей
ФК 3	Здатність розробляти нові методи інтелектуального аналізу та прогнозування в багатовимірних просторах на основі гетерогенних даних
ФК 4	Здатність використовувати метрики для оцінки точності моделей прогнозування та методів машинного і глибокого навчання на гетерогенних даних
ФК 5	Здатність розробляти високопродуктивні інформаційні технології для задач інтелектуального аналізу даних різної природи на основі хмарних інфраструктур
ФК 6	Здатність формулювати завдання та пропонувати практичні рішення щодо використання гетерогенних даних в задачах цифровізації економіки
ФК 7	Здатність до самонавчання, пошуку, аналізу, оцінки результатів експериментів з обробки гетерогенної інформації та її якості
ФК 8	Здатність керувати проектами, забезпечувати розв'язання задач цифровізації та сталого розвитку суспільства на основі гетерогенних даних
ФК 9	Здатність здійснювати науково-дослідну діяльність, приймати обґрунтовані рішення та представляти знання та ідеї щодо цифровізації економіки

Очікувані результати навчання сертифікатної програми:

ЗНАННЯ	
ЗН 1	Знати: <ul style="list-style-type: none"> - основні етапи попередньої обробки гетерогенних даних; - методи профайлінгу, трансформації, очистки та збагачення багатовимірних різнорідних даних; - статистичні та імовірнісні методи аналізу багатовимірних даних; - методи гармонізації гетерогенних даних; - методи візуалізації багатовимірних даних; - методи кореляційного та факторного аналізу гетерогенних даних.
ЗН 2	Знати: <ul style="list-style-type: none"> - основні методи кластеризації багатовимірних даних; - сучасні моделі класифікації багатовимірних даних; - сучасні алгоритми машинного навчання для гетерогенних даних; - моделі глибоких нейронних мереж для гетерогенних даних.
ЗН 3	Знати: <ul style="list-style-type: none"> - методи регресійного аналізу та прогнозування на основі багатовимірних даних різної природи; - методи еволюційної оптимізації для задач інтелектуального аналізу; - нечіткі моделі підтримки прийняття рішень на основі багатовимірних

	гетерогенних даних; - методи побудови цифрових двійників.
ЗН 4	Знати основи збору та структурування гетерогенних даних, критеріальну базу для дослідження якості даних і адекватності математичних моделей.
ЗН 5	Знати методи статистичного оцінювання параметрів багатовимірних процесів, методи оптимізації в умовах невизначеності, методи прогнозування економічних та екологічних показників на основі багатовимірних гетерогенних даних.
ЗН 6	Знати сучасні інформаційні технології зберігання та обробки великих об'ємів гетерогенних даних, технології Data Mining.
ЗН 7	Знати основи хмарних технологій, властивості та функціональні можливості сучасних хмарних платформ щодо обробки гетерогенних даних.
ЗН 8	Знати сучасні інформаційні технології для побудови цифрових двійників.
УМІННЯ	
УМ 1	Уміти: - використовувати теорію ймовірності та математичної статистики для аналізу гетерогенних даних; - розробляти моделі машинного та глибокого навчання для гетерогенних даних; - використовувати методи машинного та глибокого навчання для розв'язання прикладних задач на основі гетерогенних даних; - розв'язувати задачі регресійного аналізу та прогнозування на основі багатовимірних даних різної природи; - використовувати методи інтелектуального аналізу в задачах підтримки прийняття рішень; - будувати математичні моделі цифрових двійників для розв'язання прикладних задач бізнес-аналізу.
УМ 2	Уміти: - виконувати гармонізацію, профайлинг, трансформацію, попередню обробку, очистку та збагачення багатовимірних різнорідних даних; - оцінювати якість даних, а також планувати їх подальшу обробку для розв'язання прикладних задач; - гармонізувати та візуалізувати багатовимірні дані; - знаходити залежності, здійснювати кореляційний та факторний аналіз над великими масивами гетерогенних даних; - поєднувати дані, отримані з різних джерел, за допомогою спеціалізованого програмного забезпечення.
УМ 3	Уміти: - використовувати та налаштовувати системне та прикладне програмне забезпечення для роботи із різнорідними даними; - використовувати сучасні інформаційні технології для аналізу гетерогенних даних великого об'єму; - використовувати мову програмування Python для аналізу гетерогенних даних (Data Science); - реалізувати власні алгоритми інтелектуального аналізу на мові програмування Python; - реалізовувати методи роботи з багатовимірними масивами гетерогенних даних (Data Science) у хмарному середовищі; - розробляти інформаційні технології створення цифрових двійників, що базуються

	на використанні гетерогенної інформації.
УМ 4	Уміти: - кваліфіковано та самостійно працювати з літературою та іншими джерелами наукової інформації; - формувати опис результатів досліджень та здійснювати їх публікацію в наукових виданнях.

5. Перелік освітніх компонентів

Освітні компоненти сертифікатної програми	Кількість кредитів ЄКТС	Форма підсумкового контролю	Семестр вивчення
Методи аналізу великих гетерогенних даних	5	Іспит	10
Методи глибокого навчання на різнорідних даних	5	Іспит	10
Інформаційні технології аналізу великих гетерогенних даних	4	Залік	10
Загальний обсяг кредитів ЄКТС	14		

6. Викладання та оцінювання

Викладання та навчання	Лекції, практичні, семінарські, лабораторні заняття
Оцінювання	Види контролю результатів навчання: поточний, календарний, семестровий. Контроль проводиться згідно з Положенням про поточний, календарний та семестровий контроль результатів навчання в КПІ ім. Ігоря Сікорського Оцінювання результатів навчання здійснюється за рейтинговими системами, визначеними у силабусах навчальних дисциплін. Рейтингові системи оцінювання складені згідно з вимогами Положення про систему оцінювання результатів навчання в КПІ ім. Ігоря Сікорського

7. Ресурсне забезпечення реалізації програми

Кадрове забезпечення	Викладачі, що забезпечують викладання освітніх компонентів сертифікатної програми, є відомими фахівцями в області злиття різнорідних даних, побудови сучасних моделей штучного інтелекту на основі сучасних засобів обробки та аналізу даних. Викладачі є відомими вченими та мають високі індекси цитування у наукометричних базах даних (SCOPUS, WoS, Google Scholar)
Матеріально-технічне забезпечення	При викладанні освітніх компонентів сертифікатної програми будуть задіяні аудиторії та комп'ютерне обладнання Фізико-технічного інституту
Інформаційне та навчально-методичне забезпечення	Дисципліни забезпечуються навчальними посібниками, монографіями викладачів, які забезпечують викладання дисциплін, відеолекціями, методичними рекомендаціями до виконання лабораторних робіт, курсами Google Classroom, а також сучасними засобами комунікацій

ОПИСИ ОСВІТНІХ КОМПОНЕНТІВ

Методи аналізу великих гетерогенних даних

Курс, семестр	5 курс, весняний семестр
Обсяг	5 кредити ЄКТС (150 годин)
Мова викладання	Українська
Кафедра	Математичного моделювання та аналізу даних
Викладачі, які забезпечують викладання дисципліни	Д.т.н., професор Куссуль Н.М., д.т.н., професор Шелестов А.Ю.
Вимоги до початку вивчення (міждисциплінарні зв'язки)	Для вивчення дисципліни «Методи аналізу великих гетерогенних даних» студент має бути знайомий з основами математичного аналізу, лінійної алгебри, аналізу та підготовки даних, програмування, бажано на мові Python, мати загальне представлення про аналіз даних, математичну статистику та обчислювальні методи.
Що буде вивчатися	Основні теми, які розглядаються у курсі: <ul style="list-style-type: none"> - Основні принципи аналізу великих гетерогенних даних - Джерела гетерогенних даних та принципи їх обробки - Методи спільного аналізу та використання гетерогенних даних - Гетерогенні аналітичні системи та представлення даних.
Чому це цікаво/треба вивчати	За останні 5-7 років з'явилося багато різних джерел даних, які можна використовувати для розв'язання багатьох прикладних задач на сучасному рівні. При цьому необхідно вміти використовувати сучасні математичні моделі та методи, а також сучасне програмне забезпечення для їх реалізації та використання.
Чому можна навчитися (результати навчання)	<ul style="list-style-type: none"> - Здатність до самонавчання та продовження професійного розвитку - Вміння організації власної діяльності з освоєння наукоємних методів та технологій - Вміння здійснювати збір, опрацювання, аналіз, систематизацію науково-технічної інформації
Як можна користуватися набутими знаннями і уміннями (компетентності)	Здатність розробляти алгоритми та структури даних, програмні засоби та програмну документацію. Здатність використовувати сучасні технології програмування, математичні моделі та методи для реалізації наукоємних рішень з аналізу та обробки гетерогенних даних великого об'єму.
Заняття	Лекції, лабораторні
Інформаційне забезпечення	Навчальні посібники, відеолекції, лабораторні роботи, Google Classroom
Поточний контроль	Модульна контрольна робота, виконання лабораторних робіт, додаткові завдання
Семестровий контроль	Усний іспит або тест (при змішаній формі навчання)

Методи глибокого навчання на різномірних даних

Курс, семестр	5 курс, весняний семестр
Обсяг	5 кредити ЄКТС (150 годин)
Мова викладання	Українська
Кафедра	Математичного моделювання та аналізу даних
Викладачі, які забезпечують викладання дисципліни	Д.т.н., професор Куссуль Н.М., д.т.н., професор Шелестов А.Ю.
Вимоги до початку вивчення (міждисциплінарні зв'язки)	Для вивчення дисципліни «Методи глибокого навчання на різномірних даних» студент має бути знайомий з класичними алгоритмами лінійної алгебри та роботи з масивами даних (сортування, конкатенація, пошук елементів, обчислення статистики елементів масиву). Знати основні інструменти програмного забезпечення для обробки даних, основи Python для Data Science. Використовувати теорію математичної статистики та теорії ймовірності для розв'язання математичних задач.
Що буде вивчатися	Основні теми, які розглядаються у курсі: <ul style="list-style-type: none"> - Вступ до глибокого навчання - Основні поняття нейронних мереж - Традиційні нейронні мережі (shallow) - Глибокі нейронні мережі
Чому це цікаво/треба вивчати	Останнім часом з'явилося багато різних джерел даних, які можна використовувати для розв'язання багатьох прикладних задач на сучасному рівні. Крім цього активно розвиваються глибокі нейронні мережі, які дозволяють розв'язувати принципово інші задачі з використанням всіх наявних гетерогенних та різномірних даних. Вивченню саме таких питань і присвячена дана дисципліна.
Чому можна навчитися (результати навчання)	<ul style="list-style-type: none"> - Здатність до самонавчання та продовження професійного розвитку - Вміння організації власної діяльності з освоєння наукоємних методів та технологій - Вміння здійснювати збір, опрацювання, аналіз, систематизацію науково-технічної інформації, побудови складних інтелектуальних моделей
Як можна користуватися набутими знаннями і вміннями (компетентності)	В результаті вивчення даної дисципліни студенти зможуть: <ul style="list-style-type: none"> - використовувати теорію ймовірності та математичної статистики для аналізу гетерогенних даних; - розробляти моделі машинного та глибокого навчання для гетерогенних даних; --використовувати методи машинного та глибокого навчання для розв'язання прикладних задач на основі різномірних даних; - розв'язувати задачі регресійного аналізу та прогнозування на основі багатовимірних даних різної природи; - використовувати методи інтелектуального аналізу в задачах підтримки прийняття рішень; - будувати математичні моделі цифрових двійників для розв'язання прикладних задач бізнес-аналізу.
Заняття	Лекції, лабораторні
Інформаційне забезпечення	Навчальні посібники, відеолекції, лабораторні роботи, Google Classroom
Поточний контроль	Модульна контрольна робота, виконання лабораторних робіт, додаткові завдання
Семестровий контроль	Усний іспит або тест (при змішаній формі навчання)

Інформаційні технології аналізу великих гетерогенних даних

Курс, семестр	5 курс, весняний семестр
Обсяг	4 кредити ЄКТС (120 годин)
Мова викладання	Українська
Кафедра	Математичного моделювання та аналізу даних
Викладачі, які забезпечують викладання дисципліни	Д.т.н., професор Куссульт Н.М., д.т.н., професор Шелестов А.Ю.
Вимоги до початку вивчення (міждисциплінарні зв'язки)	Для вивчення дисципліни «Інформаційні технології аналізу великих гетерогенних даних» студент має бути знайомий з базовими принципами функціонування прикладного та системного програмного забезпечення, основні типи графічних даних та баз даних, базові знання щодо способів отримання гетерогенних даних. Мати загальне представлення щодо аналізу даних, дослідження їх статистичних властивостей, обчислювальних методів та програмування. Використовувати та налаштовувати системне та прикладне програмне забезпечення.
Що буде вивчатися	Основні теми, які розглядаються у курсі: <ul style="list-style-type: none"> - Сучасні інформаційні технології та програмне забезпечення для обробки гетерогенних даних - Основні методи обробки та аналізу гетерогенних даних - Обмін та представлення гетерогенної інформації
Чому це цікаво/треба вивчати	Останнім часом з'явилося багато різних джерел даних, які можна використовувати для розв'язання багатьох прикладних задач на сучасному рівні. Крім цього активно розвиваються глибокі нейронні мережі, які дозволяють розв'язувати принципово інші задачі з використанням всіх наявних гетерогенних та різномірних даних. Вивченню саме таких питань і присвячена дана дисципліна.
Чому можна навчитися (результати навчання)	<ul style="list-style-type: none"> - Здатність до самонавчання та продовження професійного розвитку - Вміння організації власної діяльності з освоєння наукоємних методів та технологій - Вміння здійснювати збір, опрацювання, аналіз, систематизацію науково-технічної інформації, побудови складних інтелектуальних моделей
Як можна користуватися набутими знаннями і уміннями (компетентності)	В результаті вивчення даної дисципліни студенти зможуть: <ul style="list-style-type: none"> - використовувати та налаштовувати системне та прикладне програмне забезпечення для роботи із різномірними даними; - використовувати сучасні інформаційні технології для аналізу гетерогенних даних великого об'єму; - використовувати мову програмування Python для аналізу гетерогенних даних (Data Science); - реалізувати власні алгоритми інтелектуального аналізу на мові програмування Python; - реалізувати методи роботи з багатовимірними масивами гетерогенних даних (Data Science) у хмарному середовищі;
Заняття	Лекції, лабораторні
Інформаційне забезпечення	Навчальні посібники, відеолекції, лабораторні роботи, Google Classroom
Поточний контроль	Модульна контрольна робота, виконання лабораторних робіт, додаткові завдання
Семестровий контроль	Залік або тест (при змішаній формі навчання)