

ВІДГУК

офіційного опонента, доктора технічних наук, професора, завідувача кафедри інформаційної безпеки Навчально-наукового Фізико-технічного інституту Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»,

Ланде Дмитра Володимировича

про дисертаційну роботу Манзюка Едуарда Андрійовича

**«Теоретичні та прикладні засади інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом»,**

подану на здобуття наукового ступеня доктора технічних наук

за спеціальністю 05.13.06 – інформаційні технології

## **1. Актуальність теми дисертації та зв'язок з науковими програмами планами, темами**

Інформаційні технології досягли такого рівня розвитку, який дозволяє ефективно застосовувати їх у різноманітних сферах людської діяльності. Подальший розвиток технологій супроводжується інтеграцією інтелектуальних систем. Активне впровадження інтелектуальних систем обумовлено суттєвими перевагами та новими можливостями пов'язаними з використанням інтелектуальних інформаційних технологій. Оцінка перспектив застосування інтелектуальних систем на сьогодні досягла такого рівня, що такі системи стали сферою стратегічного значення для соціального та економічного розвитку країн світу, в тому числі України. Інтелектуальні системи стали частиною цифрового порядку денного та є його ключовим компонентом. Водночас з потенційними перевагами, які передбачаються із впровадженням інтелектуальних систем, існують ризики в тому, що результати впровадження таких систем можуть суперечити принципам життя суспільства, створювати загрози безпечного

існування та загалом приводити до виникнення ймовірних катастрофічних ситуацій.

Такі загрози виникають внаслідок інтеграції інтелектуальних систем у різного роду інформаційні технології, що надає можливість машинним системам формувати рішення в області їх практичного застосування. Делегування прийняття рішень інтелектуальним системам повинно супроводжуватись впевненістю в тому, що рішення сформовані машиною будуть безпечними та не створюватимуть потенційних загроз.

На сьогодні питання недостатності довіри до інтелектуальних інформаційних технологій стало тим фактором, який визначає глибину проникнення інтелектуальних систем у різні сфери життя суспільства.

Такі обставини значною мірою вплинули на принципи розробки та впровадження інтелектуальних інформаційних технологій. Виникла потреба у розробці інтелектуальних інформаційних технологій, які б відповідали сукупності вимог довіри суспільства до таких технологій, що формує актуальну науково-прикладну проблему, для вирішення якої постає необхідність у розробці теоретичних та прикладних засад інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом.

Відомі методи розробки інтелектуальних інформаційних технологій демонструють орієнтацію, здебільшого, на технічні якісні показники їхньої ефективності. Питання довіри до інтелектуальних інформаційних технологій виходить за межі технічної ефективності та потребує нових, зокрема міждисциплінарних досліджень.

Актуальність дисертації обґрунтовується нагальною необхідністю у вирішенні завдань пов'язаних із забезпеченням довіри до інтелектуальних інформаційних технологій з метою розширення їх практичного використання у різноманітних сферах життя та діяльності суспільства, в тому числі важливих та критичних.

Дисертаційне дослідження проводилось в рамках наукової роботи кафедри комп'ютерних наук Хмельницького національного університету та

відповідно до плану держбюджетної науково-дослідної роботи Хмельницького національного університету за темою: «Розроблення інформаційної технології прийняття контрольованих людиною критично-безпекових рішень за ментально-формальними моделями машинного навчання» (номер державної реєстрації 0121U112025). В межах вказаної держбюджетної теми автором проведено обґрунтування теоретичних засад та прикладних рекомендацій щодо розробки інформаційної технології прийняття рішень у сферах важливого та критичного суспільного значення.

## **2. Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації**

Наукові положення, висновки та рекомендації, сформульовані у дисертації, є обґрунтованими, достовірними та базуються на фундаментальних наукових положеннях, мають необхідний та належний рівень узагальнень, сформовані на основі опрацювання актуальних наукових досліджень вітчизняних та закордонних науковців.

В процесі розробки інтелектуальної інформаційної технології отримання довірчих рішень використано методи аналізу та моделювання процесів, загальної теорії систем, методи онтологічного моделювання та порівняння, методи реляційної алгебри, методи класифікації та кластеризації, теорії моделювання, методи агрегації, методи теорії множин, теорії графів, методи кластерного аналізу.

Дисертаційна робота складається з анотації, вступу, семи розділів, висновків, списку використаних джерел, додатків. Основний зміст роботи викладено на 301 сторінці, загальний обсяг роботи становить 453 сторінки. Робота містить 50 рисунків, 8 таблиць та 6 додатків на 46 сторінках. Список використаних джерел нараховує 593 найменування.

У **вступі** автором представлено обґрунтування актуальності вибраного напрямку дослідження; зв'язок теми дослідження з науковими програмами, планами, темами; сформульовані мета та завдання дослідження; визначено

об'єкт та предмет дослідження; представлені методи, які було використано для вирішення поставлених завдань; сформульована наукова новизна отриманих результатів та практичне значення; подано опис особистого внеску автора у публікаціях здійснених зі співавторами; наведено опис апробації результатів проведених досліджень. Також представлено загальну структуру роботи.

У **першому** розділі «Аналіз інформаційних технологій в межах концепту довіри до рішень, прийнятих інтелектуальними системами» автором наведено зміст поняття довіри до інформаційних технологій сучасного рівня розвитку та їх вплив на розвиток промисловості та суспільства. Обґрунтовано теоретичні засади формування нового напрямку у розвитку технологій, як довірчих інтелектуальних інформаційних технологій. Проведено аналіз сучасного стану розвитку інформаційних технологій, встановлено причини та нагальну необхідність з формування довіри до інформаційних технологій у їхньому практичному застосуванні. Наведено зв'язок етичних принципів з поняттям довіри на основі людиноцентрованого підходу. Автором обґрунтовано формуванням довіри в умовах розвитку інтелектуальних інформаційних технологій в критичних та відповідальних застосуваннях. На основі проведеного дослідження автор представив обґрунтування необхідності розробки теоретичних та прикладних засад інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом.

У **другому** розділі «Формування довіри до інтелектуальних інформаційних технологій на основі етичних принципів» автором проведена розробка онтології довіри в контексті людиноцентрованого підходу. Встановлено визначальні сутності при формуванні довіри. В межах розробленої онтології визначені сутності уразливостей, загроз та викликів щодо формування довіри; високорівневі застороги необхідності довіри; заходи пом'якшення загроз щодо забезпечення довіри та зацікавлені сторони щодо наявності довіри. Подальша декомпозиція поняття довіри та заходів щодо її формування дозволила автору визначити множину етичних принципів, які у своїй сукупності дають змогу забезпечити довіру до інтелектуальних інформаційних технологій.

Для обґрунтування сукупності етичних принципів забезпечення довіри та встановлення відносної важливості кожного з розглянутих принципів, автором розроблено метод визначення відповідності складових онтології довіри до інших узагальнень етичних принципів довіри. В межах розробленого методу розглядаються синтаксичний, структурних та семантичних рівень встановлення відповідності складових довіри. Проведене автором дослідження та розроблений метод встановлення ступеня відповідності складових довіри дозволили визначити відносну важливість кожного етичного принципу та його вплив на формування довіри.

У **третьому** розділі «Методи прийняття довірчих рішень з використанням ансамблів систем прийняття рішень та сегментація даних» автором сформовані критерії вибору систем прийняття рішень для створення ансамблю. Розроблено метод формування ансамблю систем прийняття рішень, метод сегментування даних за агрегацією результатів роботи ансамблю систем прийняття рішень. Запропоновані методи дозволяють сегментувати дані за їхньою відповідністю принципам довіри та отримати високі показники якості класифікації даних за методом однокласової класифікації. Якісні показники класифікації порівнюються з відповідними показниками систем прийняття рішень, які входять до ансамблю. Проведені дослідження виявили ефективність групового підходу до класифікації даних.

У **четвертому** розділі «Метод проєкції сформованої людиною ментальної моделі прийняття рішень на машинний рівень виконання» автором розроблено метод, який дозволяє здійснити проєкцію представлення людини про роздільність даних на машинний рівень виконання. В розділі автором вирішено задачу інтеграції інтелектуальних можливостей людини у вигляді ментального представлення в системи рішень на машинному рівні виконання. Використовуються методи візуальної аналітики з метою формування ментального представлення, як системи розмежування даних. Інтеграція людини в інтелектуальну систему дозволяє підвищити рівень довіри до прийнятих рішень. У розділі автором запропоновано та розроблено метод з

побудови системи прийняття рішень, в основі якої лежить ментальне представлення сформоване людиною.

У **п'ятому** розділі «Метод аналізу складнокласифікованих даних» автором розроблено метод аналізу даних, які є сегментом загальних даних з наявними альтернативними рішеннями систем прийняття рішень в ансамблі. Такі дані визначаються як складнокласифіковані. Для такого сегменту даних автором запропоновано здійснити перерозподіл даних в гіперпросторі ознак шляхом зміни певних ознак. Ознаки, які значної мірою впливають на локальне зміщення даних, виділяються в окрему множину з встановленням міри впливу. Автором запропоновано систему метрик, згідно якої визначається міра впливу ознаки на розмежовування даних. Необхідно відзначити, що зміна ознак здійснюється в межах множин дослідження та приналежності, а також має локальний характер. Це дає змогу зберегти інформативність довірчих даних, не змінюючи їхнє просторове розташування. Важливим аспектом запропонованого підходу є забезпечення локального дослідження ознак даних поблизу зони їх розмежування згідно однокласової класифікації.

У **шостому** розділі «Інтелектуальна інформаційна технологія отримання довірчих рішень» автором запропонована інтелектуальна інформаційна технологія отримання довірчих рішень, яка базується на основі сукупності розроблених методів. У розділі автором здійснена проєкція етичних принципів забезпечення довіри до інтелектуальних інформаційних технологій в частині прийнятих рішень на практичний рівень. Розроблена сукупність методів дозволяє забезпечити реалізацію множини етичних принципів у інтелектуальній інформаційній технології в частині прийнятих рішень.

У **сьомому** розділі «Експериментальні дослідження методів сегментації даних за складовими довіри та аналізу ознак» автором проведені експериментальні дослідження ефективності розроблених методів забезпечення довіри до прийнятих рішень. Визначені кількісні показники щодо сегмента довірчих даних, якісні показники класифікації на загальному об'ємі даних. Також проведені експериментальні дослідження з аналізу впливу ознак на

розмежування складнокласифікованих даних, де представлені кількісні показники множини впливових ознак на основі розроблених метрик та множин характерного зосередження ознак. Проведені експериментальні дослідження показали ефективність запропонованої інтелектуальної інформаційної технології в частині прийняття довірчих рішень, а також продемонстрували покращення якісних показників класифікації даних.

У **висновках** автором наведені отримані в дисертаційній роботі результати досліджень, які представлені в аргументованій та логічній структурі. Науковий рівень представлених висновків та положень вказує на наукову зрілість автора та відповідає вимогам, які ставлять до докторської дисертації.

У списку використаних джерел автором у дисертаційному дослідженні використано сучасні наукової публікації, що достатньо широко охоплюють коло завдань, які вирішуються у роботі. Необхідно зазначити загальний обсяг джерел та їх належну близькість до завдань дослідження.

### **3. Достовірність наукових положень, висновків та рекомендацій**

Отримані автором наукові результати, які представлені в дисертаційному дослідженні, отримані автором особисто та є результатом наукової та аналітичної діяльності. Наукові положення, висновки та рекомендації є достовірними, що підтверджується коректністю застосування наукового підходу в межах сформульованої проблеми, відповідністю проведених досліджень меті та задачам дослідження, повнотою розв'язання поставлених задач, належним та адекватним застосуванням наукових методів дослідження.

Результати отримані автором під час виконання дисертаційного дослідження доповідались та були представлені на всеукраїнських, міжнародних науково-практичних конференціях, в тому числі тих, матеріали яких індексуються у наукометричних базах Scopus, Web of Science та складають суттєву частку від загальної кількості.

Достовірність результатів підтверджується успішним впровадженням на підприємствах ТОВ «Терралаб АЙ ТІ», ТОВ «Науково-технічна фірма

“Інфосервіс”», ТОВ «Джі Ем Хост», підприємствах-учасниках ГО «ІТ-кластер Хмельницького, ТОВ «Девлуп.про», ПП «Авіві», ПрАТ «Мікросистема», у навчальних закладах: Хмельницький національний університет, Тернопільський національний педагогічний університет імені Володимира Гнатюка, Київський національний університет імені Тараса Шевченка.

#### **4. Наукова новизна результатів дослідження**

Найбільш суттєвими результатами наукового дослідження, які отримані автором особисто та є новими у галузі інформаційних технологій сформульовано та наведено в наступному:

*вперше розроблено:*

- інтелектуальну інформаційну технологію отримання довірчих рішень, перевага якої над відомими полягає у реалізації руху інформаційних потоків формування рішень за складовими довіри, які визначені на основі запропонованої онтології довіри як сукупності етичних принципів за людиноцентрованим підходом, з виділенням сегментів інформації у вигляді типових даних з відповідністю складовим довіри та нетипових даних – як складнокласифікованих даних в умовах невизначеності прийняття рішень, і забезпеченні належних показників якості класифікації;

- метод проєкції сформованої людиною ментальної моделі прийняття рішень на машинний рівень виконання, який відрізняється від відомих тим, що з використанням методів візуального аналізу інформації здійснюється відображення результатів ментальної обробки інформації на машинний рівень інтелектуальної інформаційної системи, що дозволяє подати результати опрацювання інформації людиною у формалізованому вигляді;

- метод аналізу складнокласифікованих даних, основна перевага якого над відомими полягає у використанні локального просторового зміщення даних в області розмежування, завдяки чому з’являється можливість визначати нетипові ознаки за критерієм розмежованості даних;

*набули подальшого розвитку:*



– метод визначення нетипових ознак складнокласифікованих даних згідно з їх паралельною обробкою, який, на відміну від існуючих, виділяє обмежений, за впливом на групування даних, набір ознак у вигляді впорядкованої послідовності, що дозволило розробити метрики важливості впливу цих ознак на розмежування даних;

– метод визначення відповідності складових онтології довіри до інших узагальнень етичних принципів довіри, перевага якого над існуючими є в тому, що встановлюються структурні зв'язки між елементами порівняння та визначаються текстові описи сутностей, що дозволило використовувати в запропонованій онтології довіри сформульовані різним чином узагальнення етичних принципів довіри та визначити важливість складових онтології за описовою сутністю;

– метод групового прийняття рішень ансамблем систем прийняття рішень, який відрізняється від відомих тим, що рішення об'єднуються в групи за близькістю та формують ієрархію з узагальнення груп, завдяки чому з'являється можливість підвищити якісні показники класифікації складнокласифікованих даних порівняно зі складовими ансамблями;

*удосконалено:*

– метод формування ансамблю систем прийняття рішень, який відрізняється від відомих тим, що складові ансамблю формуються залежно від необхідного рівня інтерпретованості рішень ансамблю, завдяки чому з'являється можливість отримати рішення ансамблю з належним рівнем довіри;

– метод сегментування даних за агрегацією результатів роботи ансамблю систем прийняття рішень, перевага якого над відомими методами полягає в тому, що розв'язки складових ансамблю визначено критерієм сегментування даних, стосовно яких приймаються рішення, що уможливило сформувати сегмент даних, рішення стосовно яких відповідають складовим довіри і є одноголосними, та сегмент складнокласифікованих даних, щодо яких присутня невизначеність при прийнятті рішень.

Наукові результати отримані автором в дисертації на здобуття наукового ступеня кандидата наук у дисертаційному дослідженні поданому до захисту наукового ступеня доктора наук не використовувались.

## **5. Практичне значення результатів дослідження**

Практичне значення результатів дисертаційного дослідження визначається доведенням науково обґрунтованих положень до рівня практичних рекомендацій та застосувань, які спрямовані на підвищення довіри до інтелектуальних інформаційних технологій в частині прийнятих ними рішень на основі відповідності сукупності етичних принципів за людиноцентрованим підходом.

Розроблені методи дозволяють:

- отримати сегмент даних, рішення щодо яких вважаються довірчими за відповідністю етичним принципам в межах ансамблю систем прийняття рішень, як системи всебічного розгляду даних незалежними системами прийняття рішень;

- отримати числову оцінку якості класифікації на загальному обсязі даних за такими показниками як повнота та точність, яка за числовими показниками перевищує оцінку складових ансамблю систем прийняття рішень;

- отримати висновок щодо даних, які характеризуються наявністю альтернативних рішень за відповідними рішеннями складових ансамблю щодо даних, які не можуть вважатись довірчими;

- детектувати ознаки серед цільових груп дослідження, які мають визначальний вплив на розмежування складнокласифікованих даних, та отримати числову оцінку наявності частки таких ознак та міри їхнього впливу;

- отримати висновок щодо наявності ознак-викидів на основі розроблених метрик впливу ознак на розмежування даних з відповідною числовою оцінкою;

– сформувати ансамбль систем прийняття рішень із забезпеченням необхідного рівня інтерпретованості результатів класифікації за наявності необхідної системи прийняття рішень в ансамблі.

Практичне значення результатів дисертаційного дослідження визначається доведенням до необхідного рівня інженерних розробок з належним описом методів, отриманням висновків та числових результатів.

Підтвердженням практичного значення отриманих результатів є їх впровадження у практичній діяльності підприємств, що підтверджено відповідними документами, які наведені у дисертаційній роботі.

Результати дисертаційного дослідження рекомендуються до впровадження на підприємствах при розробці інтелектуальних інформаційних технологій для застосування в системах важливого та критичного значення.

## **6. Дискусійні положення та зауваження**

Даючи позитивну оцінку дисертаційному дослідженню загалом, змістовності та обґрунтованості наукових положень, висновків та рекомендацій, необхідно зазначити ряд дискусійних положень та зауважень:

1. В дисертаційній роботі автором проведено дослідження поняття довіри до інтелектуальних інформаційних технологій, здійснення декомпозиція цього поняття на складові за етичними принципами, та в подальшому визначається відносна вагова частка кожної складової довіри. Надалі розглядається довіра до інтелектуальних інформаційних технологій в частині прийнятих ними рішень, що є тільки частиною загального поняття. Оскільки інтелектуальні системи розглядаються як системи прийняття рішень, очевидно, що довіра в частині прийнятих рішень має вагому частку в загальному понятті. Однак автором не наведена оцінка вмісту розглянутого аспекту в загальному понятті, хоча складові довіри розглянуті та в межах проведених досліджень така можливість наявна, що варто було зробити.

2. З множини вхідних даних виділяється сегмент даних, рішення щодо яких можна вважати довірчими за сукупністю відповідності етичним принципам, та сегмент даних, рішенням щодо яких не можна довіряти. Для забезпечення необхідного рівня інтерпретованості до ансамблю долучається відповідний класифікатор, як необхідний. Для класифікації недовірчих даних з ансамблю визначаються класифікатори, які в поєднанні рішень максимізують необхідні якісні показники. Однак в цей набір не завжди долучаються класифікатори необхідного рівня інтерпретованості, відповідно отримуються дані з різним рівнем інтерпретованості, про що в роботі варто було б зазначити.

3. На рисунку 7.9 (стр. 322) продемонстровано ліс дерев траєкторій зменшення міжкласових ребер. Зображенню слід було б надати більше інформативності та провести більш детальний аналіз отриманих результатів. Очевидно, що не всі гілки дерев лісу ведуть до максимального покращення розмежування класів. Додатковий аналіз покращив би дисертаційне дослідження.

4. Автором на сторінці 242 описуються три типи ознак, які можна виділити за своїм характерним впливом на групування та розмежування даних. Далі досліджуються тільки ознаки, зміна яких призводить до покращення розмежування. Інші типи ознак не досліджуються. Варто було б провести повний аналіз ознак, в межах виділених груп, з метою цілісного розгляду характеристик ознак, для отримання повної картини, а не зосереджуватись на одній групі.

5. На рисунку 6.3, сторінка 289 наведено схему отримання рішень з вибором оптимального набору з ансамблю систем прийняття рішень за цільовим показником якості. Однак дана схема описує формування набору класифікаторів для частини даних, які автором визначені як складнокласифіковані. Варто було б навести схематичне зображення з отримання рішень для всього обсягу даних, а надалі, послідовно розкрити зміст з отримання рішень для складових сегментів даних. Це є важливим аспектом, не наведеним автором у роботі.

6. В роботі в розділі 3 та в розділі 7 проводяться дослідження з розподілу показників кореляції рішень класифікаторів та формування кінцевого рішення ансамблем класифікаторів. Слід було б детально навести відмінності у проведених дослідженнях за результативністю та цілями їх проведення, що за наявного опису є недостатнім.

7. З використанням однокласової класифікації в роботі виділяється сегмент довірчих даних. Далі аналіз ознак здійснюється для сегмента даних, які не відносяться до довірчих. Однак, якщо в роботі дослідження ведеться з точки зору довіри до рішень, які приймаються ансамблем, для повноти втілення цього поняття слід було провести аналіз ознак, які є об'єднуючими та характерними для основного класу.

Зазначені зауваження мають дискусійний характер, не знижують наукове значення дослідження та не впливають на загальну позитивну оцінку роботи.

### **Загальний висновок**

Дисертаційна робота Манзюка Е.А. «Теоретичні та прикладні засади інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом» є самостійно виконаним, завершеним науковим дослідженням, присвячена вирішенню актуальної проблеми забезпечення довіри до інтелектуальних інформаційних технологій в частині прийнятих ними рішень, має вагоме теоретичне та практичне значення. За рівнем представлення наукових та прикладних результатів, логіки викладення, застосування наукових методів дослідження дисертаційна робота відповідає вимогам, які ставляться до докторських дисертацій.

Зміст дисертаційної роботи повною мірою відповідає темі дослідження, отримані результати мають теоретичне обґрунтування, практично підтвердженні та відповідають спеціальності 05.13.06 – інформаційні технології.

Таким чином, приймаючи до уваги рівень наукового обґрунтування результатів дисертаційного дослідження, експериментальну перевірку, прикладне значення, рівень висвітлення у наукових публікаціях, вважаю, що

дисертаційна робота відповідає вимогам пп. 6, 7, 8, 9 «Порядку присудження та позбавлення наукового ступеня доктора наук» затвердженого Постановою Кабінету Міністрів України № 1197 від 17 липня 2021 р., а її автор, Манзюк Едуард Андрійович, заслуговує на присудження наукового ступеня доктора технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент

доктор технічних наук, професор, завідувач  
кафедри інформаційної безпеки Навчально-  
наукового Фізико-технічного інституту

Національного технічного університету  
України «Київський політехнічний інститут  
імені Ігоря Сікорського»



Дмитро ЛАНДЕ

Підпис професора Ланде Д.В засвідчую

Вчений секретар

КПІ ім. Ігоря Сікорського



Валерія ХОЛЯВКО