

Хмельницький національний університет

Міністерство освіти і науки України

МАНЗЮК ЕДУАРД АНДРІЙОВИЧ



УДК 004.896:004.912:004.048

**ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ ЗАСАДИ ІНТЕЛЕКТУАЛЬНОЇ
ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ОТРИМАННЯ ДОВІРЧИХ РІШЕНЬ ЗА
ЛЮДИНОЦЕНТРОВАНИМ ПІДХОДОМ**

05.13.06 – Інформаційні технології

(12 – інформаційні технології)

Реферат
дисертації на здобуття наукового ступеня
доктора технічних наук

Хмельницький – 2022

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Обґрунтування вибору теми дослідження. Стрімкий розвиток та широке розповсюдження інформаційних технологій зумовлені перспективами практичного застосування обчислювальних систем. Зокрема, запровадження інтелектуальних інформаційних технологій значною мірою розширило можливості людини для прийняття рішень як через розв'язання складних обчислювальних задач, так і через автономне прийняття рішень. Розвиток інтелектуальних інформаційних технологій досягає того рівня, при якому їм можна делегувати завдання щодо прийняття рішень і при подальшому розвитку створювати умови формування суб'єктності у прийнятті рішень. Прикладами такого застосування є автоматичні системи торгівлі, системи діагностування захворювань, безпілотні автомобілі та автомобілі з автопілотами, системи керування літальними апаратами різного призначення, в тому числі для сільського господарства чи військового призначення, системи розпізнавання об'єктів або ідентифікації супротивника тощо.

Однак, надаючи інтелектуальним інформаційним системам можливість автономно приймати рішення, виникає проблема відповідальності за прийнятті рішення. Технічні системи, на жаль, можуть мати у своїй роботі, під впливом різних причин, технічні та програмні несправності. Неправильно прийняте рішення автопілотом автомобіля є загрозою життю пасажирів, а у випадку прийняття суспільно критичних рішень, неправильне рішення має ще більш значні наслідки. Виникає проблема безпечного інтегрування інтелектуальних інформаційних систем в життя суспільства.

Таким чином, рівень розвитку інтелектуальних інформаційних технологій та переваги їх широкого застосування зумовлюють появу нових завдань, які виникають за умов формування нової соціально-технологічної системи. В поточний час інтелектуальна технічна система виходить на рівень партнерства з людиною. Розвиток інтелектуальних інформаційних технологій приводить до того, що людина існує поряд з інтелектуальною технічною системою як автономним суб'єктом прийняття рішень, і таке існування повинно мати партнерський характер. Фундаментом існування партнерства щодо інтелектуальних інформаційних технологій та делегування можливості прийняття рішень є довіра до усіх проявів та складових інтелектуальних інформаційних технологій. Тому, враховуючи актуальність та важливість відносин людини та інтелектуальної технічної системи, виникає необхідність у розробленні методів побудови систем прийняття рішень, в основу яких покладено складові поняття довіри до інтелектуальних інформаційних технологій.

Проблемі розвитку інтелектуальних інформаційних технологій з позиції взаємодії з людиною та суспільством присвячено широкий спектр як українських, так і іноземних робіт та наукових шкіл: Лепрі Б. (Lepri B.), Алам К. (Alam K), Ламмарш Т. (Lammarsch T), Бахманн Н. (Bachmann N), Брем А. (Brem A.), Ванг Дж. (Wang J.), Шнайдер Т. (Schneider T.), Чен М. (Chen M.), Чіатто Г. (Ciatto G.), Гербер Л. (Gerber L.), Донг Х (Dong X), Дуббер М.Д. (Dubber M.D.), Ендерт А. (Endert A.), Лютге К. (Lütge C.), Джобін А. (Jobin A.), Кейм Д. (Keim D.), Квон Б. (Kwon B.), Мюллер В.К. (Müller V.C.), Остерлінк Д. (Oosterlinck D.), Тібес С.

(Thiebes S.), Пападакіс Г. (Papadakis G.), Перера П. (Perera P.), Рен Д. (Ren D.), Діттеріх Т.Г. (Dietterich T.G.), Шіпп К.А. (Shipp C.A.), Бріньолфссон Е. (Brynjolfsson E.), Ключин Д.А., Ляшко С.І., Ланде Д.В., Гуляницький Л.Ф., Яковлев С.В., Крак Ю.В., Глибовець А.М., Бодяньський Є.В., Анісімов А.В., Говорущенко Т.О., Литвиненко В.І., Теслюк В.М., Бармак О.В., Савенко О.С., Пелешко Д.Д., Шаховська Н.Б., Глибовець М.М., Лисенко С.М., Березький О.М., Турбал Ю.В., Федоров Є.Є., Куусуль Н.М., Бідюк П.І. та ін.

Актуальність проблеми взаємодії людини та інтелектуальних інформаційних технологій підтверджується приділеною значною увагою наукової спільноти, яка зумовлена потребою у вирішенні пов'язаних нагальних проблем, які донедавна не виникали, а наразі уповільнюють прогресивний розвиток суспільства загалом. Однак, існуючі наукові роботи не розглядають побудову практичних рішень інтелектуальних інформаційних технологій в частині методів побудови систем прийняття рішень, в основу яких покладено забезпечення довіри за її складовими до інтелектуальних інформаційних технологій.

Таким чином, актуальною **науково-практичною проблемою** є недостатній рівень довіри до інтелектуальних інформаційних технологій в частині прийняття ними рішень на практичному рівні.

Актуальність вирішення науково-практичної проблеми полягає у наявності протиріччя між зростаючим розповсюдженням застосування інтелектуальних інформаційних технологій у різних сферах людського життя, включаючи сфери критичного значення, та наявністю засторог до їх застосування, причиною яких є недостатній рівень довіри до рішень, які генерують такі системи.

Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційна робота виконана відповідно до планів науково-дослідних робіт Хмельницького національного університету за держбюджетною темою «Розроблення інформаційної технології прийняття контрольованих людиною критично-безпекових рішень за ментально-формальними моделями машинного навчання» за номером державної реєстрації 0121U112025. У вказаній науково-дослідній роботі здобувач брав участь як відповідальний виконавець у розробці принципів, методів та інформаційних технологій прийняття важливих та відповідальних рішень інтелектуальними системами прийняття рішень з використанням формального та ментального представлення.

Мета і завдання дослідження. Метою дисертаційної роботи є забезпечення довіри до інтелектуальних інформаційних технологій шляхом розроблення теоретичних та прикладних засад інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом, яка забезпечує довіру до рішень отриманих інтелектуальними інформаційними системами відповідно до сукупності етичних принципів.

Для досягнення поставленої мети дисертаційної роботи необхідно виконати ряд завдань:

- провести аналіз поняття довіри до інтелектуальних інформаційних технологій як концепту людиноцентрованого підходу розвитку етичних інтелектуальних технічних систем;

- розробити структуру концепту довіри до інтелектуальних інформаційних технологій та провести декомпозицію на структурні елементи;
- розробити метод проєкції сформованої людиною ментальної моделі прийняття рішень на машинний рівень виконання;
- розробити метод визначення відповідності складових онтології довіри до інших узагальнень етичних принципів довіри;
- розробити метод формування ансамблю систем прийняття рішень;
- розробити метод групового прийняття рішень ансамблем систем прийняття рішень;
- розробити метод сегментування даних за агрегацією результатів роботи ансамблю систем прийняття рішень;
- розробити метод аналізу складнокласифікованих даних з визначенням нетипових ознак та викидів;
- розробити метод визначення нетипових ознак складнокласифікованих даних згідно з їх паралельною обробкою;
- розробити інтелектуальну інформаційну технологію отримання довірчих рішень за людиноцентрованим підходом;
- здійснити експериментальні дослідження методів сегментації даних за складовими довіри та аналізу ознак методів синтезу систем прийняття довірчих рішень та аналізу ознак даних.

Об’єкт дослідження – процес забезпечення довіри до інтелектуальних інформаційних технологій.

Предмет дослідження – методи та засоби інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом.

Методи дослідження. В межах дисертаційної роботи визначено ряд завдань, для вирішення яких було використано методи аналізу та моделювання процесів, теорії систем, методи онтологічного моделювання та порівняння, методи реляційної алгебри для моделювання довіри до інтелектуальних інформаційних технологій, методи класифікації та кластеризації, теорії моделювання для розробки моделей групового отримання рішень з використанням методів агрегації корельованих рішень та формування областей даних відповідно до показників довіри, методи теорії множин, теорії графів, методи кластерного аналізу для розробки методу аналізу ознак даних відповідно до виявлення викидів на характерних множинах щодо приналежності ознак і формування гіперпросторового групування даних з мінімізацією втрати інформативності.

Наукова новизна одержаних результатів. Наукові результати, які визначають новизну дисертаційної роботи полягають у наступному:

вперше розроблено:

1) 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;

2) 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;

3) 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:

4) 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;

5) 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;

6) 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 із складовими ансамблями;

удосконалено:

7) 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;

8) 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.

Практичне значення отриманих результатів полягає в тому, що запропоновані методи отримання рішень інтелектуальною інформаційною системою за показниками довіри та якісними показниками набули практичної реалізації, що дозволило отримувати множини довірчих рішень. Застосування ансамблевого

підходу дозволило отримувати необхідне дискримінантне розмежування рішень за цільовими показниками якості. Стало можливим отримання наборів ознак для слабороздільних даних, які визначають максимально можливу роздільність, та встановити міру впливу ознак на розмежованість даних для ознак певної цільової групи.

Результати проведених практичних досліджень дозволили визначити складові концепту довіри з проєкцією на програмний рівень побудови систем прийняття рішень. Практичне застосування розроблених методів побудови систем прийняття рішень дозволило виділити сегмент даних, рішення щодо яких характеризуються високим ступенем довіри, та сегмент нетипових даних поблизу зони розмежування. Поряд з цим, вдалося отримати високі показники якості розмежування складнороздільних даних та запропонувати застосування методів для аналізу ознак з виділенням викидів.

Розроблені методи можуть бути застосовні в межах широкого спектра задач, які належать до сфери інтелектуальних інформаційних систем загалом, оскільки дозволяють отримати результати з високими показниками якості. Особливе значення має практичне застосування розроблених методів систем прийняття рішень в інтелектуальних інформаційних технологіях для сфер важливого та критичного значення, які потребують отримання рішень з високими показниками довіри. Сфери безпекових, соціально важливих, військових, фінансових, критичних рішень для суспільства та життя людей із застосуванням інтелектуальних інформаційних технологій, а також такі, що можна до них віднести, є тими предметними областями, де можуть бути використані розроблені інтелектуальні інформаційні системи.

Реалізація результатів та впровадження. Результати дисертаційної роботи впроваджено у навчальному процесі Хмельницького національного університету, Київського національного університету імені Тараса Шевченка, Тернопільського національного педагогічного університету імені Володимира Гнатюка. Розроблену інтелектуальну інформаційну технологію отримання довірчих рішень, яка базується на розроблених методах реалізації етичних принципів, що визначені за людиноцентрованим підходом, впроваджено: для ПЗ систем контролю рішень автоматичних інтелектуальних систем на підприємстві ТОВ «Терралаб АЙ ТІ», де вона дала можливість отримати сегмент довірчої інформації 79 %, забезпечити якісні показники повноти 96 %, точності 86 %, визначити сукупність нетипових ознак на рівні 2.5 %; для ПЗ в сервісах отримання довірчих рішень для систем пошуку інформації на підприємстві ТОВ «Науково-технічна фірма “Інфосервіс”», де вона дала можливість отримати сегмент довірчої інформації 86 %, визначити частку найбільш впливових ознак близько 3.5 %, підвищити розмежування складнокласифікованої інформації на 63 %; для ПЗ систем класифікації потокової інформації на підприємстві ТОВ «Джі Ем Хост», де вона дала можливість отримати сегмент довірчої інформації 83 %, підвищити класифікацію вхідної інформації за точністю на 1.3 %, повнотою 3.8 %, отримати множину найбільш впливових ознак на розмежованість інформації близько 3 %; для ПЗ з використанням інтелектуальних інформаційних систем обробки потоків інформації на підприємствах-учасниках ГО «ІТ-кластер Хмельницького», де вона дала можливість отримати сегмент довірчої інформації близько 78 %, визначити частку найбільш впливових ознак, яка становить від 1.5 % до 4.2 % відносно загальної кількості

ознак; для ПЗ в системі обробки інформації на підприємстві ТОВ «Девлуп.про», де вона дала можливість отримати сегмент довірчої інформації 86 %, підвищити точність класифікації на 1.7 % та повноти на 4.2 %; для ПЗ при обробці потоку інформації на ПП «Авіві», де вона дала можливість отримати сегмент довірчої інформації близько 84 %, визначити частку найбільш впливових ознак на розмежування інформації, яка становить близько 3 %; для ПЗ при обробці інформації в ПрАТ «Мікросистема», де вона дозволила отримати сегмент довірчої інформації в межах 76 %, отримати показники класифікації інформації за повнотою 97 %, точністю 95 %, а також множину найбільш впливових ознак близько 3 %.

Особистий внесок здобувача. Усі наукові результати дисертаційної роботи отримані автором особисто. Роботи [5, 9, 12] виконувались без співавторів. У публікаціях написаних в співавторстві, автору належать такі результати: [13, 15, 37] – інформаційна технологія класифікації даних з використанням метода редукції ознак з багатомірного простору; [17, 38] – метод класифікації даних на основі агрегації рішень ансамблю систем прийняття рішень; [48] – метод агрегування рішень для послідовних даних; [21, 40] – метод формування ансамблю систем прийняття рішень за групуванням рішень для забезпечення необхідних якісних показників; [51, 53] – метод системи прийняття рішень на основі довіри до інтелектуальних інформаційних технологій; [26, 47] – методи побудови систем прийняття рішень для простору з багатьма ознаками; [41, 49] – метод класифікації даних із застосування методів візуального аналізу; [58] – метод проєкції ментального представлення на машинний рівень формального представлення; [43, 54, 57] – експериментальне дослідження формування ансамблів систем прийняття рішень на основі групованості рішень; [45] – підхід з використання групованості даних для детектування нетипових даних; [23, 39] – метод визначення нетипових ознак даних в приграничних даних з використання дерев кортежів ознак; [55, 59] – підхід з моделювання групованості даних за приналежністю до класів; [42, 50] – дослідження використання візуалізації даних для побудови ментального представлення для класифікації; [31] – метод структурного порівняння відповідності складових концепту довіри та структурованого домену; [29, 32] – підхід з визначення нетипових ознак при аналізі даних; [30] – формальна структура довіри до інтелектуальних інформаційних технологій; [27, 33] – підхід з класифікації даних на основі візуальної аналітики за просторовою групованістю; [34] – визначення характеристики для формування ансамблю систем прийняття рішень; [28] – метод формування групованостей рішень з використанням ансамблів; [14, 18] – підхід з розробки людино та машиноцентрованого представлення за областю формування; [19, 36] – підхід з формування ментального представлення при класифікації даних на основі візуальної аналітики; [60] – дослідження ефективності застосування метода ментально-формальної проєкції моделі; [61] – дослідження з формування ментальної моделі за візуальним поданням даних; [62] – дослідження з використання систем прийняття рішень; [22, 3, 1] – дослідження ефективності застосування ментального представлення при класифікації даних; [16, 2, 4] – розробка критерії формування систем прийняття рішень в ансамбль; [7, 8, 46, 52, 56, 10, 11, 44] – обґрунтування та аналіз застосування інтелектуальних інформаційних технологій в системах отримання важливих та довірчих рішень; [24, 25, 35] – підхід з формування технології класифікації даних з кусково-лінійними методами розмежування; [6, 20] – дослідження групування даних за сукупністю ознак при класифікації.

Апробація результатів дисертації. Основні результати дисертаційної роботи доповідались та обговорювались на міжнародних та всеукраїнських конференціях та форумах: International Workshop of IT-professionals on Artificial Intelligence (ProfIT AI 2021) September 20-21, 2021 (Kharkiv); II International Scientific Symposium "Intelligent Solutions" (IntSol-2021) September 28-30, 2021 (Kyiv - Uzhhorod); Fourth International Workshop on Computer Modeling and Intelligent Systems (CMIS-2021) April 27, 2021 (Zaporizhzhia); 1st, 2nd, 3rd International Workshop on Intelligent Information Technologies & Systems of Information Security (IntelITSIS-2020) June 10-12, 2020, (IntelITSIS-2021) March 24–26, 2021, (IntelITSIS-2022) May 25-27, 2022 (Khmelnyskyi); 9th International Conference "Information Control Systems & Technologies" (ICST-2020) September 24–26, 2020 (Odessa); 11th, 12th International Scientific and Practical Conference of Programming (UkrPROG 2018) May 22–24, 2018, (UkrPROG 2020) September 15-16, 2020 (Kyiv); 14th International conference on pattern recognition and information processing (PRIP'2019) May 21-23, 2019 (Minsk, Belarus); IEEE 14th, 15th International Conference on Computer Sciences and Information Technologies (CSIT-2019) September 17-20, 2019 (Lviv), (CSIT-2020) September 23-26, 2020 (Zbarazh); IEEE International Conference on Advanced Trends in Information Theory (ATIT-2019) December 18-20, 2019, (ATIT-2021) December 15-18, 2021 (Kyiv); 44th WILGA 2019 Symposium on Photonics Applications and Web Engineering (Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments) May 25 – June 2, 2019 (Wilga, Poland); International Scientific Conference "Intellectual Systems of Decision-making and Problems of Computational Intelligence" (ISDMCI-2019) May 21-25, 2019, (ISDMCI-2020) May 25-29, 2020, (ISDMCI-2021) May 26-28, 2021 (Kherson - Zalizny Port); 2nd International Conference On Intelligent Computing & Optimization (ICO-2019) October 3-4, 2019 (Koh Samui, Thailand); IV международная научно-практическая конференция "Информатика и прикладная математика" Сентябрь 25-29, 2019 (Алматы, Казахстан); XXXIII International Conference «Problems of decision making under uncertainties» (PDMU-2019) January 24 – February 1, 2019 (Hurgada, Egypt); XI, XII, XIII Всеукраїнська науково-практична конференція «Актуальні проблеми комп'ютерних наук» (АПКН-2019) 14-15 листопада 2019, (АПКН-2020) 9-10 листопада 2020, (АПКН-2010) 15-16 жовтня 2021 (Хмельницький); Всеукраїнська науково-практична конференція «Інтелектуальний потенціал – 2018» 14-16 листопада 2018 (Хмельницький); Міжнародна конференція молодих науковців «Сучасні технології в механіці» 19-21 квітня 2018 (Хмельницький); XI, X Міжнародна науково-практична конференція «Інформаційні управляючі системи та технології» (ІУСТ–2020) 24–26 вересня 2020, (ІУСТ–2021) 23 - 25 вересня 2021 (Одеса); XVIII Міжнародна науково-технічна конференція «Штучний інтелект та інтелектуальні системи» («Artificial Intelligence and Intelligent Systems», AIPS'2018) 18 – 19 жовтня 2018 (Київ); Науково-практична конференція «Актуальні проблеми теорії керуючих систем у комп'ютерних науках» (АПТКС-2021) 21-24 грудня 2021 (Слов'янськ); XVII, XVIII, XIX міжнародна науково-практична конференція «Математичне та програмне забезпечення інтелектуальних систем» (МПЗІС-2019) 20-22 листопада 2019, (МПЗІС-2020) 18-20 листопада 2020, (МПЗІС-2021) 17-19 листопада 2021 (Дніпро); VIII, IX, X Всеукраїнська науково-практична конференція «Глушковські читання - Ідеї

академіка В.М. Глушкова і сучасні проблеми штучного інтелекту», 29 листопада 2019, «Глушковські читання - Ідеї академіка В.М. Глушкова і сучасні проблеми теоретичної кібернетики», 18 грудня 2020, «Глушковські читання - Історія, сучасний стан та тенденції цифрового розвитку суспільства», 2 грудня 2021 (Київ); International scientific conference «IX Ukrainian-Polish scientific dialogues» October 20-23, 2021 (Khmelnyskyi – Kamianets-Podilskyi); International Workshop Advances & Challenges in Computing (A2C) October 20, 2021 (Ternopil); щорічних наукових конференціях професорсько-викладацького складу Хмельницького національного університету.

Публікації. За темою дисертації з викладенням основних її результатів опубліковано 62 наукових праці, з них: 12 статей у наукових виданнях, включених до Переліку наукових фахових видань України, 5 статей у наукових періодичних виданнях інших держав, індексованих у наукометричних базах Scopus та/або Web of Science Core Collection, віднесених до третього квартилю (Q3) або до другого квартилю (Q2) відповідно до класифікації SCImago Journal and Country Rank, що згідно із Наказом МОН України №1220 від 23.09.2019 р. прирівнюються до двох або трьох публікацій відповідно; 2 розділи у колективних монографіях, індексованих у наукометричній базі Scopus, 40 публікацій, які засвідчують апробацію матеріалів дисертації (у тому числі 16 індексованих у наукометричних базах Scopus та/або Web of Science); 3 свідоцтва про реєстрацію авторського права на твір.

Структура та обсяг дисертації. Дисертаційна робота складається з анотації, вступу, семи розділів, висновків, списку використаних джерел із 593 найменувань на 76 сторінках та 6 додатків на 46 сторінках. Загальний обсяг дисертації становить 453 сторінки, з них 301 сторінка основного тексту, 50 рисунків, 8 таблиць.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обгрунтовано актуальність обраної теми дисертаційної роботи, вказано зв'язок роботи з науковими програмами, планами та темами, сформульовано мету та основні завдання дослідження. Визначено об'єкт, предмет та методи дослідження, вказана наукова новизна та практичне значення отриманих результатів роботи. Надано характеристику особистого внеску здобувача у роботах, виконаних у співавторстві, також представлені відомості про апробацію результатів роботи. Подано опис структури та обсягу дисертації.

У **першому розділі** проведено дослідження сучасного стану розвитку інформаційних технологій із використанням інтелектуальних інформаційних складових. Проаналізовано вплив, застосування та інтегрування інтелектуальних інформаційних технологій (ІІТ) у сфері життя як окремої людини, так і суспільства загалом. Встановлено, що ІІТ проявляються не тільки у вигляді технологічних рішень, а також є визначальним фактором розвитку та трансформації суспільства. Таким чином, інтеграцію ІІТ слід розглядати як соціотехнологічну систему.

Проведений аналіз дозволив визначити, що особливу увагу слід приділяти не тільки удосконаленню та розвитку ІІТ за показниками ефективності функціонування та виконання поставлених задач, а також впливу ІІТ на людину та суспільство. Крім того, приділяється значна увага проблемі інтеграції ІІТ у життя та розвиток

соціальних інституцій, формувань та системи в цілому. З цією метою були проведені ґрунтовні дослідження з визначення викликів використання ІТ. Аналіз показав, що актуальність, нагальність та важливість вирішення питань широкого та повсюдного застосування ІТ проявляється в зростанні суспільного інтересу, діяльності професійних організацій, державних та міжнародних структур, корпоративного сектору. Інтеграція ІТ у суспільне життя супроводжується впровадженням державних регуляцій, розробленням нормативних документів, стандартів, рекомендації та інше.

Було встановлено, що соціальні принципи на основі людиноцентрованого підходу реалізуються в суспільстві, яке за організацією готове до проактивного соціального впровадження ІТ. Відповідно, такі принципи повинні бути реалізовані в практичних дослідженнях та застосовуватись в проєктуванні ІТ. Розглянуті підходи показують, що в соціотехнологічному суспільстві необхідно забезпечити реалізацію таких принципів: орієнтація на людину із людиноцентрованим підходом; принцип отримання суспільством пізнання та етики ІТ з рівним та постійним доступом; принцип захисту конфіденційності та приватності; забезпечення безпеки з належною оцінкою ризиків; формування чесного конкурентного середовища; принцип справедливості, підзвітності, прозорості систем ІТ; принцип постійної та стійкої інновації. Таким чином, поєднання цих принципів в єдину систему з підпорядкуванням фундаментальним цінностям дозволяє створити середовище для суспільного розвитку та трансформації з врахуванням громадських та індивідуальних інтересів для ефективного використання ІТ. Визначено, що поняття “довіри до ІТ ” відіграє ключову роль за можливості формування обставин в яких існує невизначеність, складність взаємозв’язків з ймовірним небажаним розвитком подій та їх послідовностей, а також є ключовим фактором, який базується на етичних принципах і визначає розвиток суспільства як соціотехнологічної системи.

Проведені дослідження показали:

- ІТ на сучасному рівні розвитку здійснюють визначальний вплив на соціальні відносини на структурному рівні, що формує нові соціальні відносини та визначає нову форму середовища розвитку особистості та суспільства;

- відомий цілий ряд національних та міжнародних програм розвитку та впровадження ІТ, дорожніх карт інтеграції ІТ, нормативних документів, стандартів, професійних рекомендацій, які підпорядковані спільній меті формування безпечного та ефективного середовища застосування ІТ, однак це стосується визначення соціальних базових принципів, проте відсутні методи та механізми імплементації вимог на технологічному та технічному рівні;

- проєктування, розробка, впровадження ІТ є динамічною та міждисциплінарною областю дослідження, яка потребує нових підходів, що виходять за межі технічних та технологічних рішень та потребують використання останніх наукових здобутків з філософії, соціології, психології, в яких ІТ необхідно розглядати як соціотехнологічну систему.

Аналіз системного підходу з розвитку ІТ дозволив встановити, що:

- відомі методи проєктування і розробки розглядають ІТ на основі технологічних підходів як технічного напрямку, окремим напрямком в межах гуманістичного підходу є дослідження вимог щодо ІТ як соціально відповідальних

технологій, проте відсутні розробки з єдиним соціотехнологічним підходом, який представлений методами імплементації рішень базових принципів людиноцентрованого підходу;

– відсутні методи розробки ІТ на основі соціотехнологічного підходу, які орієнтовані на інтеграцію людини та її інтелектуальних можливостей в ІТ на рівні розробки методів та технічних рішень і забезпечення реалізації соціальних принципів;

– визначені та сформовані принципи, на основі яких повинні бути реалізовані ІТ, не доведені до рівня практичної реалізації з декомпозицією складових структурних компонентів в частині вимог та відповідними технологічними механізмами забезпечення втілення в конкретних рішеннях.

Наприкінці розділу, на основі проведених досліджень та аналізу, сформульовану науково-прикладну проблему, яка полягає у підвищенні рівня довіри до інтелектуальних інформаційних технологій в частині прийняття ними рішень на практичному рівні. Поставлено мету та визначено ряд завдань, які необхідно вирішити в межах наукового дослідження. В наступних розділах подано методи та практичні рішення для вирішення представленої проблематики дослідження.

У **другому розділі** розроблено онтологію стандартизованого поняття довіри до ІТ. Це дало можливість визначити основні концепти, які дозволяють сформулювати положення довіри, є змістовною частиною поняття довіри до ІТ, зумовлюють необхідність її існування та являють собою для неї загрозу. На базі онтології предметної області проведена декомпозиція структурних змістовних понять. В подальшому визначені характеристики концепту формування довіри.

Центральним концептом стандарту є довіра до ІТ, яка визначається як здатність ІТ відповідати очікуваним сподіванням стейкхолдерів, піддається перевірці та має можливості до реалізації такої перевірки.

Онтологія створена на базі стандарту та визначена групами характеристик, які в сукупності формують довіру до рішень, що приймаються з використанням підходів ІТ. Стандарт ISO / IEC TR 24028 “Інформаційні технології – Інтелектуальні інформаційні системи” (Огляд питань довіри до інтелектуальних інформаційних систем) описується в межах понять відповідної предметної області. Стандарт визначає високорівневі концепти, які формують поняття довіри. Також визначається множина та поняття стейкхолдерів, як основних зацікавлених сторін в понятті довіри. Потенційні вразливості ІТ та загрози, які пов’язані з ними визначені в концептах вразливості, загроз та викликів. Для задоволення потреби стейкхолдерів в довірі встановлюється концепт заходів пом’якшення наслідків – можливі засоби контролю, рекомендації та настанови, які можуть знизити вплив відомих уразливостей ІТ.

Структура предметної області стандарту «Інформаційні технології – Інтелектуальні інформаційні системи» на основі онтологій представлено у вигляді:

(1)

де *Con* - множина концептів визначених в межах стандарту;

Rel - множина відношень $SbArTrw = \langle Con, Rel \rangle$ між концептами.

Множина концептів формується на основі інформації стандарту $Con = \{\{Vul, Thr, Chal\}, \{MtMs\}, \{HLConc\}, \{Sth\}, \{TrW\}\}$ та встановлюється множина відношень між ним $Rel = \{"reduce", "weaken", "form", "define", "need"\}$.

Встановлено, що концепт “заходи пом’якшення наслідків” ($MtMs$) формує довіру до ІТ. Відповідно, складові елементи цього концепту формують множину характеристик та визначають функціональний зв’язок щодо довіри $TrW = f(MtMs)$. Довіра (TrW) відповідно до стандарту є функцією основних категорій (фразових категорій). Структура “заходи пом’якшення наслідків” ($MtMs$) має наступну форму:

$$MtMs = \{Ct, subCt, Ch\}, \quad (2)$$

де Ch - характеристики, $subCt$ - підкатегорії, Ct - категорії.

Елементи множини сформовано в ієрархії $Ch \subseteq subCt \subseteq Ct$ та пов’язані функціональним залежностями з атрибутами довіри $\langle Ct, subCt, Ch \rangle = f(\{atr_i\}_{i=1}^n)$. Кожна категорія є множиною характеристик, які об’єднані стандартом відповідно до семантичного значення та утворюють семантичні кластери.

Враховуючи ієрархічні структури формування концептів в стандарті, найбільш зручним встановлено отримання вищої ієрархії концептів, які представлено у вигляді однорідної структури фразових категорій. Аналіз показав, що в межах стандарту це рівень підкатегорій. Перелік підкатегорій концепту $MtMs$: прозорість, пояснювальність, контрольованість, стратегії зменшення упередженості, конфіденційність, достовірність, резильєнтність, робастність, пом’якшення несправностей апаратного забезпечення системи, функціональна безпека, тестування, аналізованість, використання, застосовність. Множина ‘заходів пом’якшення’ наслідків на рівні підкатегорій має вигляд $MtMs = \{mtms_i\}_{i=1}^n, n=14$. Тут n - кількість фразових підкатегорій концепту. Множина відношень між складовими концепту $MtMs$ складається з відношення “*isPartOf*”, тобто $Rel_{MtMs} = \{"isPartOf"\}$.

Концепт $MtMs$ представлено у вигляді:

$$M_{MtMs} = \left\{ \begin{array}{l} Trs, Exp, Cont, Bsred, Prv, Rlb, Rsl, Rbs, Flmit, \\ Sffun, Tst, Evl, Us, Apl, "isPartOf" \end{array} \right\}, \quad (3)$$

де Trs - прозорість, Exp - пояснювальність, $Cont$ - контрольованість, $Bsred$ - стратегії зменшення упередженості, Prv - конфіденційність, Rlb - достовірність, Rsl - резильєнтність, Rbs - робастність, $Flmit$ - пом’якшення несправностей апаратного забезпечення системи, $Sffun$ - функціональна безпека, Tst - тестування, Evl - аналізованість, Us - використання, Apl - застосовність.

З метою перевірки правильності вибраного підходу щодо визначення формоутворюючих складових концепту довіри до ІТ було розроблено визначення відповідності складових онтології довіри до інших узагальнень етичних принципів довіри.

Порівнюють усі сутності онтології з усіма сутностями представленими в структурованому домені. Порівнювання відбувається згідно з припущенням

закритого світу, відповідно до якого, якщо будь-які пари сутностей, що не згадані у порівнянні, вважаються такими що не мають зв'язків.

Для порівняння використано комбінацію трьох різних підходів оцінки подібності сутностей: синтаксичний, структурний та семантичний (рис. 1.). Порівняльна оцінка дається виходячи з того наскільки подібними є сутності, і розглядається як міра ступеня заміщення одного класу іншим з можливістю використання.

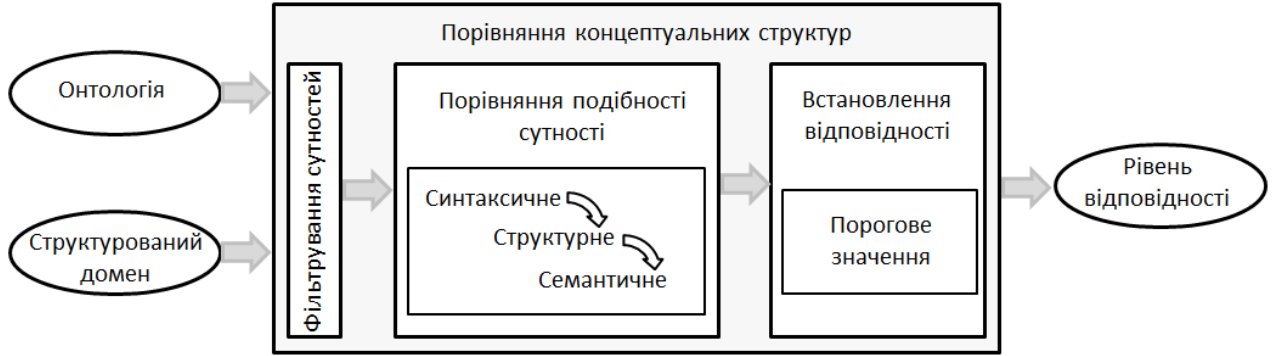


Рис. 1. Загальна структура системи порівняння концептуальних структур онтології та структурованого домену.

В межах домену дослідження представляється у вигляді ідентифікатора сутності $id \in I$, позначення $lb \in L$ та опису $dsc \in Dsc$. Зміст сутності описується кортежем $ent = \langle id, lb, dsc \rangle$. Два елемента змісту сутності ent_1 та ent_2 збігаються, якщо вони співвідносяться з однією і тією ж сутністю предметної області з об'єктивної реальності. Ці дві сутності є взаємозамінними тобто справедливе співвідношення $ent_1 \equiv ent_2$, або в межах дослідження $ent_o \equiv ent_{Ds}$. В загальній множині сутностей $\forall ent_o \in Ent_o, \forall ent_{Ds} \in Ent_{Ds}, Ent_o \cup Ent_{Ds} = Ent$, яка є об'єднанням сутностей онтології та структурованого домену, відповідно наявні повтори сутностей.

Ефективність порівняння концептуальних структур стосовно онтології визначається співвідношенням, яке встановлює наскільки розроблена онтологія відповідає стану предметної області: $Ef_o = \frac{|Link|}{|Ent_o|}$.

Фільтрацію сутностей множин Ent_o та Ent_{Ds} здійснено через функцію подібності $Ent_o \times Ent_{Ds} \xrightarrow{f_{sim}} \mathbb{R}$, з визначеним пороговим значенням trh_{sim} . Відповідно, отримано множину подібності на рівні зв'язків

$$Link_{O,filter} = \left\{ \begin{array}{l} link_{ent_{O,i}, ent_{Ds,j}} \left| \left(ent_{O,i}, ent_{Ds,j} \right) \in Ent_o \times Ent_{Ds}, ent_{O,i}, \right. \\ \left. i \in |Ent_o|, j \in |Ent_{Ds}|, f_{sim} \left(ent_{O,i}, ent_{Ds,j} \right) \geq trh_{sim} \right. \end{array} \right\}.$$

Оскільки досліджується відповідність онтології O стосовно Ent_{Ds} , множина зв'язків визначається з базовим відношенням стосовно сутностей, які визначені в онтології $Link_{O,Ds,filter} = Ent_o \times_{trh_{sim}} Ent_{Ds}$.

Множина подібності сутностей на рівні зв'язків формується за інтуїтивним припущенням про відповідність вибраної сутності онтології певним сутностям структурованого домену.

Формування множини зв'язків $Link_{O,D_s,filter}$ здійснюється згідно з такими етапами:

1. Для кожної сутності $ent_{O,i}$, $i \in |Ent_O|$ множини сутностей онтології Ent_O вибирається сукупність сутностей структурованого домену Ent_{D_s} . Відкидаються сутності істино негативні $Ent_O \times_{TN} Ent_{D_s}$.

2. Наступним кроком є зменшення кількості зв'язків множини $Link_{O,D_s,filter}$ до знайдених ймовірних відповідностей з базовою множиною Ent_O . Зменшення $Link_{O,D_s,filter}$ здійснюється згідно з пороговим значенням подібності $Link_{O,D_s,filter} = \sigma_{\geq th_{sim}}(Link_{O,D_s,filter})$.

3. На цьому кроці відбувається перехід до кроку 1 та мінімізується кількість сутностей множини онтології Ent_O , які не мають зв'язків $\min(Ent_O \triangleright_{TN} Ent_{D_s}) = \min(Ent_O - Ent_O \times_{TN} Ent_{D_s})$. Ця умова є важливою, оскільки на цьому кроці визначається кількість сутностей Ent_O для яких не було знайдено порівнювальних концептуальних структур в Ent_{D_s} .

Онтологію порівняння концептуальних структур онтології довіри та структурованого домену представлено у такому вигляді:

$$O_{match} = \langle Ent_O, Ent_{D_s}, Link_{O,D_s,filter}, \rho, \sigma_{trh_{match}} \rangle. \quad (4)$$

Функція відбору $\sigma_{trh_{match}}$ визначає обмеження множини $Link_{O,D_s,filter}$ використовуючи встановлення відповідностей зв'язків базової множини Ent_O на основі порогового значення ступеня відповідності trh_{match} . Тип відповідності $\rho = \{\equiv, \sqsubseteq, \sqsupseteq\}$. Міра ступеня відповідності отримує базову сутність з Ent_O та присвоює кожному елементу підмножини зв'язків з $Link_{O,D_s,filter}$, в яких присутня базова сутність, відповідне число в межах $[0, \dots, 1]$. Таким чином, крайні значення визначають у випадку рівності 1 - ідентичність сутностей, або у випадку рівності 0 - відсутність значимого зв'язку між цими сутностями, тобто визначають силу зв'язку.

Відповідність для синтаксичного рівня порівняння визначено таким чином:

$$\begin{aligned} & syntactic(ent, ent') = \\ & \frac{w_{label} \sigma(matchClass(ent, ent')) + \sum w_{prop_vector} (matchProp(ent, ent'))}{|prop_vector|} \\ & \frac{w_{label} + \sum w_{prop_vector}}{|prop_vector|} \end{aligned} \quad (5)$$

$$s.t. \ syntactic(ent, ent') \neq syntactic(ent', ent);$$

$$ent \in Ent_O, ent' \in Ent_{D_s}, syntactic(ent, ent') \in [0, \dots, 1].$$

Тут ent та ent' сутності множин порівняння, $\sigma(x) = \frac{1}{1 + e^{5(x-0.5)}}$ функція сигмоїда, w_{label} та w_{prop_vector} вагові коефіцієнти класу та властивостей класу.

Відповідність для структурного рівня порівняння:

$$structure(ent, ent') = \begin{cases} \frac{|A| + |B|}{2|\pi_{ent}(Link_O)|}; \\ A = \sigma_{link_{Ds}(ent', y) \exists link_{O, Ds, filter}(x, y) \wedge \exists link_{O}(ent, x) \wedge (\forall x \in Ent_O) \wedge (\forall y \in Ent_{Ds})}(\pi_{ent'}(Link_{Ds})); \\ B = \sigma_{link_{O, Ds, filter}(x, y) \exists link_{O}(ent, x) \wedge (\forall y \in Ent_{Ds}) \wedge (\forall z \in Ent_O)}(Link_{O, Ds, filter}), \end{cases} \quad (6)$$

$$s.t. \quad structure(ent, ent') \neq structure(ent', ent), structure(ent, ent') \in [0, \dots, 1],$$

$$\exists link_{O, Ds, filter}(ent, ent'), ent \in Ent_O, ent' \in Ent_{Ds}, \exists ! link = \sigma_{\max_w}(Link), w \in [0, \dots, 1].$$

Структурне порівняння формується з використанням зв'язків. Досліджується порівняння сутності $ent \in Ent_O$ із наявними зв'язками на множині $Ent_O - Link_{O(ent, x)}$, де $x \in \{x | link_{O(ent, x)}\}$. Тобто отримується множина зв'язків сутності порівняння з базової множини порівняння онтології довіри з іншим сутностями з цієї ж множини. Сутність дослідження подібна сутності або підмножині сутностей з множини Ent_{Ds} . Це визначає функцією відбору зв'язків між множинами Ent_O та Ent_{Ds} - $\sigma_{ent}(Link_{O, Ds, filter}) \neq \{\}$. Якщо множина наявних сутностей $\sigma_{ent}(Link_{O, Ds, filter}) > 1$ досліджуються усі подібності на наявних зв'язках з сутністю ent на множині $Link_{O, Ds, filter}$.

З метою отримання сукупності знань в межах певної концепції, яка представляється у вигляді іменованої сутності вводиться більш загальне поняття, яке узагальнює сутності на метарівні, та представляє певну еталонну сутність стосовно до завдань порівняння. Відповідно позначається метарівень сутностей - Інформаційна Рамка *ScIn* (Score Information). Призначення інформаційних рамок узагальнити знання, які представлені у вигляд набору неоднорідних сутностей, і є відповідним семантичним представлення одного і того елемента навколишнього світу.

Формальне представлення інформаційної рамки визначається таким чином:

$$ScIn = \langle Ent, Prop(Des), Scp \rangle. \quad (7)$$

Тут Ent - множина іменованих сутностей; $Prop(Des)$ - множина властивостей, які формують іменовані сутності та визначаються множиною описів Des ; Scp - рамки властивостей визначаються кортежем.

Відповідність для семантичного порівняння сформовано такою залежністю:

$$semantic(ent, ent' \Rightarrow scin) = \frac{\left| \sigma_{prop_{Ds} \in Prop_{Ds} \wedge (Fun_{semanticPropertie}(prop_O, prop_{Ds}) > threshold)}(\pi_{ent}(Prop_O)) \right|}{|\pi_{ent}(Prop_O)|}. \quad (8)$$

Семантичну відповідність визначено з використанням семантичної відповідності властивостей та є функцією від сукупності описів dsc , які отримуються з домену онтології $dom O$ та структурованого домену корпусу $dom Ds$:

$$semanticPropertie(prop_O, prop_{Ds}) = f(\{dsc_O\}, \{dsc_{Ds}\}). \quad (9)$$

В дослідженні оцінюється кількість знайдених відповідних властивостей на множині властивостей онтології відносно сутності дослідження $\pi_{ent}(Prop_O)$ та обмежених функцією відбору з тих властивостей, які мають відповідність на множині Ds , тобто визначені $semanticPropertie(prop_O, prop_{Ds})$, значення якої більше порогового $threshold$.

На основі розробленого метода було визначено важливість етичних принципів відповідності в межах структурованого домену (рис. 2.).

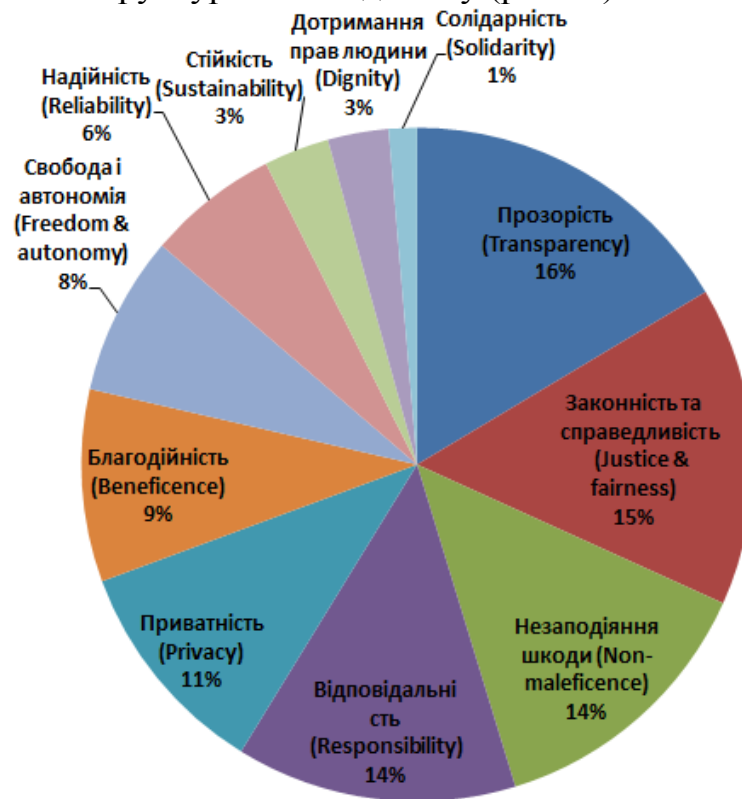


Рис. 2. Розподіл етичних принципів ІТ за відносною важливістю принципу на основі структурованого домену

Застосування розробленого метода встановлення відповідності показало ефективність запропонованого підходу з формування змістовності концепту довіри до ІТ виходячи з принципів та рекомендацій етики ІТ. Представлені концепти формування довіри до ІТ, підтверджуються контент-аналізом корпусу. Усі концепти онтології формування довіри до ІТ знайшли відповідне підтвердження на рівні головних етичних принципів.

У **третьому розділі** розроблено методи прийняття довірчих рішень з використанням ансамблів систем прийняття рішень та сегментації даних.

Відмінність кожної системи прийняття рішень, яка в межах дослідження визначається як класифікатор, встановлено як відношення даних правильно класифікованим цим класифікатором до даних, які правильно класифіковані також і іншими класифікаторами відповідно до рівня рангу. Співвідношення визначено в межах множини класифікатора. Результати обрахунку можуть змінюватися в широких межах. Параметр точність визначено як відношення множини неправильно класифікованих даних до універсальної множини даних. Цей параметр має відносний характер з максимальним значенням $p < 1$. Співвідношення (12) призначене для підсилення відмінності у відповідності до точності та дозволяє поєднати в одному параметрі ці два значимих фактори.

Алгоритм 1.

Input: індексація розмічених даних U

Set: r

forall $c \in C$ **do**

$$t \leftarrow \frac{\left| \left\{ x^y \mid x^y \in \bigcup \bigcap C_c^r \setminus C_c \right\} \right|}{\left| \left\{ x^y \mid x^y \in \bigcup \bigcap C_c^r \right\} \right|}; \quad (10)$$

$$p \leftarrow \frac{|\{C_c \setminus U\}|}{|\{U\}|}; \quad (11)$$

$$d_{c(r)} \leftarrow t/p. \quad (12)$$

end forall

sort $\{d_{c(r)}\}$

return $d_{c(r)}$

Таким чином, визначається така характеристика рішенням в ансамблі як відмінність за параметром точності. Відмінність формується за встановленим рангом r . Вона дозволяє врахувати як ранг відмінності рішення, так і точність. Подвійна назва повною мірою визначає призначення характеристики. Отримано таку характеристику класифікатора як *відмінність рангу r за параметром точності*. Ця характеристика призначена для використання в ансамблі та може бути використана як в безпосередньому вигляді (10) – відмінність, так і за параметром точності (12). Може бути використаний будь-який інший параметр, який застосовний для характеристик рішень та змінюється в межах (0, 1).

Коефіцієнт подібності визначається за мінімальною спільністю перетинів множин правильно класифікованих даних. Такий підхід використовується для порівняння якості класифікатора при зміні його параметрів класифікації, що дозволяє оцінювати якісний вплив параметрів на класифікатори.

Розглянуто опис коефіцієнта подібності k_s для використання в комплексах алгоритмів з метою формування набору алгоритмів $\{c\}$, які мають максимальну відмінність з огляду на множини правильно класифікованих даних $X \times Y$. Елементи

$X \times Y$ визначаються як $x^y = (x, y)$. Для встановлення набору класифікаторів та коефіцієнта подібності запропоновано алгоритм, який на універсальній множині U описує перетини правильно класифікованих даних з використанням трансцендентного індексування елементів U .

Алгоритм 2.

Input: Індксація даних для тестування U

$S_{\cup} \leftarrow 0$

$S_{\cap} \leftarrow 0$

forall the $x^y \in U$ **do**

for $c \in C$ **do**

$k_c \leftarrow |\{x^y \mid x^y \in c\}|$ (13)

end for

end forall

sort $\{k_c\}$

for k_c **from** $\max k_c$ **do**

if $!\exists \{x^y \mid x^y \in c\} \in C_c^1$ **then** (14)

delete c **from** $\{c\}$

$k_c = 0$

end if

end for

sort $\{k_c\}$

while $k_c \neq 0$ **from** $\max k_c$ **do**

$S_{\cap} \leftarrow S_{\cap} \cup \left\{ x^y \mid x^y \in \bigcap^c C_c^2 \setminus \bigcap^c C_c^{>2} \right\}$ (15)

$S_{\cup} \leftarrow S_{\cup} \cup \left\{ x^y \mid x^y \in \bigcup_{n=1}^c C_n \right\}$

end while

$k_s \leftarrow \frac{|S_{\cap}|}{|S_{\cup}|}$

return $k_s, \{c\}$

На початковому етапі проіндексовано простір U . Далі обраховано коефіцієнт кожного з класифікаторів як кардинальне число множини правильно класифікованих даних класифікатором (13). Оскільки належним є використання класифікаторів, що правильно класифікують найбільшу кількість даних (класифікатор має найбільший коефіцієнт), відповідно сортовано класифікатори за цим показником.

Застосування алгоритму дозволяє з набору класифікаторів отримати комплекс, який характеризується максимальною ефективністю з класифікаторів, що розглядаються, та розрахувати коефіцієнт порівняльної оцінки як коефіцієнт

подібності сукупності комплексу класифікаторів. Наступним етапом є видалення класифікаторів, які не мають безпосереднього перетину з універсальним простором. Таким чином, визначаються класифікатори в яких множинна правильних відповідей поглинається множиною іншого класифікатора або сукупністю множин. Тобто інформаційна цінність таких класифікаторів є незначною в множині класифікаторів, які розглядаються, за тієї обставини, що існують класифікатори або їх об'єднання, які узагальнюють тобто поглинають класифікуючі властивості таких класифікаторів (14). Відповідно, такі класифікатори не несуть інформаційної цінності і їхнім коефіцієнтам присвоюється нуль, що вказує на те, що їх використовувати недоцільно. Починаючи з найбільш ефективного класифікатора, визначено перетини множин правильних класифікацій. Якщо виявляються області в яких при наступному класифікаторі перетинаються більше ніж множини двох класифікаторів, такі області не враховуються, так як відбувається дублювання класифікуючих перетинів (15).

Метод формування множини систем прийняття рішень визначається такими кроками:

- формування пулу потенційних класифікаторів з належними якісними характеристики стосовно класифікації на множині вхідних даних;
- визначення параметрів відмінності класифікаторів;
- визначення коефіцієнта подібності за мінімальною спільністю перетинів множин правильно класифікованих даних;
- відбір та формування множини класифікаторів з максимальними якісними показниками та мінімальними параметрами подібності.

Граничний випадок використання ансамблів з простим голосуванням. При задачі бінарної класифікації, де $\{f_i(x)\}_{i=1}^M$ - бінарні класифікатори. Для випадкових величин, які задаються відповідним ймовірнісним простором $\xi_i = I[f_i(x)] \sim Ber(p)$, $p(0,1)$. При цьому $E[\xi_i] = p$, $D[\xi_i] = p(1-p)$. Необхідною умовою є незалежність класифікаторів, тоді $\{\xi_i\}_{i=1}^M$ незалежні. Використано $\bar{\xi} = \frac{1}{M} \sum_{i=1}^M \xi_i$ та $E[\bar{\xi}] = p$, $D[\bar{\xi}] = \frac{p(1-p)}{M}$. Так як ансамбль $F(x)$, який складається з $\{f(x)_1, \dots, f(x)_M\}$ голосуванням більшістю голосів, неправильно класифікує при умові $\bar{\xi} > 0.5$.

Ймовірність помилки:

$$P(F(x)) = \sum_{i=\frac{M}{2}+1}^M \frac{M!}{(M-i)!i!} p^i (1-p)^{(M-i)}. \quad (16)$$

Для граничного випадку за умови $M \rightarrow \infty$ значення $D[\bar{\xi}] \rightarrow 0$ розподіл величини $\bar{\xi}$ вироджується. При збільшені величини M ймовірність помилки ансамблю $P(F(x))$ наближається до нуля, якщо $p < 0.5$. На підставі цього при побудові ансамблю простим голосуванням можна говорити про можливість побудови ідеального класифікатора.

Відповідно, рішення ансамблю буде мати підвищену точність, оскільки неправильні рішення класифікатора компенсуються відповідними правильними рішеннями інших класифікаторів. Щоб отримати очікуваний результат класифікатори повинні бути різноманітними.

Умови формування ансамблю класифікаторів з множини можливих:

$$C = \{c | A(c)\}. \quad (17)$$

$$A(c_i) = \begin{cases} c_i = \arg \max_{j \in [1, \dots, W]} b_j, b_j = \begin{cases} \bigcap_{k=1}^i C_k; \\ \bigcup_{k=1}^i C_k \setminus U. \end{cases} \\ \bigcup_C C \setminus U \neq \{ \}. \end{cases} \quad (18)$$

Далі формується поєднання рішень систем прийняття рішень (СПР) в групуваності.

Метод формування ансамблю систем прийняття рішень складається з наступних кроків:

- визначаються якісні показники класифікації на вхідних даних та множині відібраних класифікаторів;
- визначається необхідний рівень інтерпретованості класифікації з необхідним обов'язковим залученням класифікатора до ансамблю;
- формується ансамбль класифікаторів з долученням до ансамблю класифікаторів, використанням методу формування множини систем прийняття рішень у формі відношення потенційних класифікаторів та наявного ансамблю.

Присутність сильної позитивної кореляції між окремими рішеннями СПР всередині ансамблю сприяє їхній агрегації в групи. Рішення кожної СПР групи сильно співвідноситься з рішеннями інших членів цієї групи. Рішення групи представляє рішення утворюючих його рішень СПР, а кожна індивідуальна СПР делегує своє рішення групі. Подальша наступна агрегація рішень груп формує рішення ансамблю. Такий процес делегування рішення рівню вищого порядку дозволяє створити ієрархічну структуру формування рішення ансамблю. Формування груп дозволяє поетапно зменшити дисперсію в ансамблі. Відбувається заміна дисперсії групи на зміщення вищого ієрархічного порядку. Ієрархічна структура ансамблю формує умови поділу рішень по рівнях локальності, а процес делегування поетапно глобалізує локальні рішення. Важливий наслідок такого процесу полягає в тому, що глобальні рішення можуть на локальному рівні бути неправильними, а також локальні рішення бути відмінними від глобального.

Метод сегментування даних за агрегацією результатів роботи ансамблю систем прийняття рішень складається з таких кроків:

- застосування класифікаторів ансамблю стосовно вхідних даних;
- формування множини даних, відносно яких консолідовані рішення класифікаторів ансамблю;
- формування множини даних, відносно яких наявні альтернативні рішення серед класифікаторів ансамблю.

Головною умовою мінімальності помилки передбачення ансамблю є рівність дисперсії груп ансамблю СПР і протилежність їх рішень з врахуванням агрегативної функції:

$$\begin{cases} \sigma_i^2 = \sigma_k^2, i \neq k \in \{1, 2\}; \\ \text{sign}P(C_{L_i}(x)) \neq \text{sign}P(C_{L_k}(x)), x \in X, i \neq k \in \{1, 2\}. \end{cases} \quad (19)$$

Система співвідноситься з незміщеною дисперсією результатів ансамблю СПР на рівні глобального оцінювання набору даних. Симетричність дисперсії забезпечує доповнюваність рішень груп СПР при використанні їх в ансамблях при незміщеній оцінці та окремих СПР у загальному випадку.

Проведено аналіз роботи оцінок класифікаторів на підставі визначення інформаційної цінності ознак класифікації. Розглянуто етапи, які проводяться при попередній обробці документів та класифікації документів на базі сучасних підходів. Класифікація проводиться з використанням як окремих класифікаторів, так і компонентного підходу ансамблю класифікаторів. Ансамблі класифікаторів використовують композицію зі зміною навчальної вибірки та поєднанням класифікаторів. Подано підхід, який дозволяє визначити елементи класифікації, які несуть інформаційну цінність та володіють узагальнюючими властивостями шляхом оцінки міри подібності між ними на підставі результатів класифікації.

Запропонована ієрархічна структура ансамблю СПР з використанням групування для класифікації даних підтверджує основні переваги використання прийняття групових рішень. При цьому, не порушуючи загальної структури і значень параметрів прогнозування, запропонований метод дозволяє отримати необхідні поєднання СПР.

У **четвертому розділі** запропоновано метод проєкції сформованої людиною ментальної моделі прийняття рішень на машинний рівень виконання. Для цього було встановлено цілі використання робочого процесу візуальної аналітики, які орієнтуються на одержання СПР: людиноцентрований робочий процес візуального аналізу будує ментальне представлення, машиноцентрований робочий процес візуального аналізу – будує формальне представлення. Представлення розглядається як інформаційний процесор і механізм прийняття рішень. Формальне та ментальне представлення відрізняються тільки споживачами, якими є машина або людина. Запропоновано дві концепції конструювання представлення системи прийняття рішень на базі синхронізації представлень і використовуючи протилежне представлення. Застосовуючи концепцію 'використання протилежного представлення' розроблена інформаційна технологія, яка дозволяє машині одержати представлення на базі ментального представлення та демонструє можливість використання цієї концепції. Це дозволяє машині повною мірою використовувати інтелектуальні можливості людини на базі представлення.

Запропоновано інформаційну технологію, яка дозволяє максимізувати інформаційне наповнення простору ознак об'єктів. Для цього зменшують розмірність у просторі можливих ознак графічного візуального зображення. Подальше зменшення розміру може виявити приховану структуру даних і дозволяє знайти приховані особливості. Основним фактором інформаційних технологій є мінімізація втрат інформаційних даних та візуальне графічне навчальне

представлення управління класифікацією інформації. Інформаційна технологія, побудована із застосуванням запропонованого методу, забезпечує гнучкий інструмент класифікації даних із можливістю вказувати межі класів.

Використання цього підходу дозволяє:

1. провести візуалізацію даних об'єктів в редукованому просторі узагальнених ознак з метою оцінки їх розподілу;
2. провести аналіз отриманих групованостей об'єктів з метою визначення характеру сукупностей і розподілів;
3. максимально ефективно оцінити розташування і обриси ліній, які необхідно задати для поділу сукупності об'єктів на необхідні класи;
4. використовуючи задані лінії обмеження класів будувати гіперплощини-аналоги, які визначають межі класів у вхідному мірному просторі;
5. використовувати зазначену межу простору для подальшої класифікації.

Інформаційна технологія складається з таких кроків:

1. формування матриці попарних відстаней на базі вхідних даних;
2. знаходження квадрата відстаней матриці відстаней;
3. використання подвійного центрування матриці;
4. визначення власних значень і власних векторів матриці;
5. оптимізація карти алгоритмом SMACOF.

Формування правил нелінійної класифікації проводиться такою послідовністю дій:

1. формування кусково-лінійних візуальних обмежень класу в редукованому просторі;
2. розрахунок опорних точок-правил для класу;
3. трансформація точок-правил в багатовимірний простір;
4. побудова гіперплощин в багатовимірному просторі на базі трансформованих точок;
5. формування правил для класу в багатовимірному просторі з використанням обмежувальних гіперплощин.

Криві задаються у вигляді кусково-лінійної структури з різним необхідним ступенем дискретизації, утворюючи криві в першому наближенні. В результаті при додаванні нового об'єкта можна чітко вказати, до якого класу він належить. На початковому етапі проходить навчання, в якому вказуються області та їхню приналежність до певного класу. Всі об'єкти конкретної області відносно меж мають певне просторове положення. Це положення формує правила приналежності об'єкта до класу. Після навчання при класифікації нового об'єкта визначається набір правил. Об'єкт належить до того класу, з наборами правил об'єктів якого він збігається. Набори правил визначають набір ознак приналежності об'єкта до класу. Ці набори ознак використовуються для побудови дерева рішень.

Інформаційна технологія має обмеження. Візуально представлені дані повинні бути візуально роздільні та згруповані. Якщо людина не побудує ментальне представлення, класифікувати дані буде неможливо. Інформаційна технологія була розроблена для демонстрації концепції 'використання протилежного представлення'. Формальне представлення повністю базується на ментальному представленні. Можливості машини при класифікації даних не використовувалися.

Пропонується метод, який дозволяє максимізувати інформаційну складову сукупності ознак об'єктів простору. Для цього використовується зменшення простору ознак на простір можливого графічного візуального представлення. Крім того, зниження розмірності дозволяє розкрити приховану структуру даних і дозволяє знайти латентні особливості. Це дає змогу максимально інформативно представити дані для аналізу співвідношення взаємозв'язків ознак. На основі аналізу графічно визначаються нелінійні межі класів з проектуванням їх в початковий гіперпростір. Таким чином відбувається формування СПР класифікації.

У **п'ятому розділі** розроблено метод аналізу складнокласифікованих даних. Нетипові дані, які належать одному класу та мають незначні за величиною відмінності в ознаках, базуючись на просторовому розташуванні, мають більшу просторову щільність за критерієм компактності. Таким чином, збільшення компактності даних класу збільшує їхню подібність, що відповідно зменшує просторові відстані. Відповідно, це дозволило визначити викиди ознак в граничних даних однокласової класифікації на основі мінімізації двочасткового графа мінімального остового дерева.

Дані у багатовимірному просторі представлено у вигляді графа.

Відстань між вершинами графа G :

$$\|E\|_1 = \sum_{u,v \in V(G)} dist_G(u,v) \text{ s.t. } diag(A) = 0, \quad (20)$$

де $dist_G(u,v)$ відстань між вершинами u та v в графі G , відповідно довжина ребра графу e , $|e| = dist_G(u,v)$. $\|\cdot\|_1 \in l_1$ обмеження норми.

Тоді за визначенням щільність даних $Dens$:

$$max(Dens_{V_{Class}}) \stackrel{def}{=} min \|E_{V_{Class}}\|_1. \quad (21)$$

Мінімізація відстаней нормованого простору даних класу призводить до побудови мінімального остового дерева на елементах класу:

$$lim_{\lambda \rightarrow a} (min \|E_{V_{Class}}\|_1) \rightarrow \|E_{V_{Class}(MST)}\|_1. \quad (22)$$

Відповідно:

$$max(Dens_{V_{Class}}) \stackrel{def}{=} min \|E_{V_{Class}(MST)}\|_1. \quad (23)$$

Позначивши λ як параметр зміни ознак, a - граничне значення параметра змін.

Побудова мінімального остового дерева (МОД) здійснюється на усіх нетипових даних. Із МОД можна виділити три типи субграфів. Субграф з ребрами, які з'єднують вершини, що являються даними, які належать класу $subMST_{Class}(V,E) = \{v | \forall v \in D_{Class}\}$. Субграф з ребрами, які поєднують дані поза межами класу $subMST_{NonClass}(V,E) = \{v | \forall v \in D_{NonClass}\}$. Субграф з ребрами, які поєднують вершини, що належать даним з цих двох попередніх множин $subMST_{Class-NonClass}(V,E) = \left\{ e \left| e_{x_i, x_j} \in E, x \in D_{Trans}, i \in [1..|V_{Class}|], j \in [1..|V_{Non-class}|] \right. \right\}$.

Субграф утворюється з мінімального остового дерева та будується на вершинах

підмножин $V = V_{Class} \cup V_{NonClass}$, де загальна кількість вершин тотожна нетиповим даним $V \equiv D_{Trans}$.

Визначено множину ребер МОД, які з'єднують вершини класу і ребра та не є інцидентними до вершин, які не належать класу $\left\{ e \mid e_{x_i, x_j} \in E, x \in V_{Class}, x \notin V_{Non-class}, i, j \in [1..|V_{Class}|] \right\}$.

МОД визначається наступним чином:

$$MST(G(V)) = \bigcup_e \left\{ e_{u,v} \in E \mid \min \sum d_G(u, v), \forall u, \forall v \in V \right\}. \quad (24)$$

На множині V_{Class} необхідно мінімізувати відстані для субграфів з ребрами, які є інцидентними вершинам класу.

За цієї умови МОД, яке побудоване на усіх даних, розбивається на субдерева, і в сукупності представляє ліс, оскільки МОД будується на усіх даних $MST(V)$, а компактність даних визначається для вершин класу субграфів МОД:

$$FMST(V_{Class}) = \bigcup_{\forall e \in E(u, e), \forall u \in V_{Class}} subMST(V) \text{ s.t. } subMST(V) \subset MST(V). \quad (25)$$

$FMST(V_{Class})$ є поєднанням частин МОД.

У найкращому випадку ліс субграфів на вершинах класу ідентичний одному субграфу МОД тобто:

$$\exists! subMST(V_{Class}) \left(\left(\left| \left\{ e_{u,v} \mid \forall u \in V_{Class}, \forall v \in V_{Non-class} \right\} \right| = 1 \right) \wedge \right. \\ \left. subMST(V_{Class}) \subset MST(V) \right). \quad (26)$$

Для забезпечення розмежування необхідно локально змінювати положення окремих елементів даних максимально зберігаючи загальне просторове положення та форму групи класу.

Відповідно, найкраща форма розмежування даних при існуванні єдиних субграфів, які являють собою дерева на відповідних даних:

$$MST(V) = \exists! subMST(V_{Class}) \cup \exists! subMST(V_{Non-class}) \cup \dots \quad (27)$$

$$\text{s.t. } subMST(V_{Class}) \& subMST(V_{Non-class}) \subset MST(V), V = V_{Class} \cup V_{Non-class}.$$

$MST(V)$ у виразі (27) в загальному випадку є лісом, оскільки в такому представленні $|E(MST)| \neq |V(MST)| - 1$, що є необхідним для існування дерева. Необхідне подальше об'єднання дерев лісу.

Аналіз ребер МОД показав:

$$E_v(MST) = \overbrace{\left\{ e_{u,v} \mid (\forall u, \forall v) \in V_{Class} \right\}}^{E_{V_{Class}}} \cup \overbrace{\left\{ e_{u,v} \mid \forall u \in V_{Class}, \forall v \in V_{Non-class} \right\}}^{E_{V_{Class} \cdot V_{Non-class}}} \cup \overbrace{\left\{ e_{u,v} \mid (\forall u, \forall v) \in V_{Non-class} \right\}}^{E_{V_{Non-class}}}. \quad (28)$$

Відповідно, до послідовності викладення (28), виділені в окремі множини ребра, вершини яких інцидентні вершинам класу V_{Class} , ребра, які з'єднують вершини не одного типу по відношенню до класу та ребра з інцидентними вершинами поза межами класу $V_{Non-class}$.

В граничному стані:

$$\lim_{FMST \rightarrow \exists! subMST} \left(\left| \left\{ e_{u,v} \mid \forall u \in V_{Class}, \forall v \in V_{Non-class} \right\} \right| \right) = 1, \quad (29)$$

де

$$FMST \rightarrow \exists! subMST : \begin{cases} FMST(V_{Class}) \rightarrow \exists! subMST(V_{Class}); \\ FMST(V_{Non-class}) \rightarrow \exists! subMST(V_{Non-class}). \end{cases} \quad (30)$$

В кінцевому представленні $MST(V)$ є деревом та забезпечується умова $|E(MST)| = |V(MST)| - 1$.

$$MST(V) = \exists! subMST(V_{Class})$$

$$\bigcup E(B)$$

$$\bigcup \exists! subMST(V_{Non-class})$$

(31)

$$s.t. (subMST(V_{Class}), B, subMST(V_{Non-class})) \subset MST(V), V = V_{Class} \cup V_{Non-class}.$$

Множина ребер двочасткового графа B :

$$E(B) = E(MST(V)) \setminus \left(\left\{ e_{u,v} \mid (\forall u, \forall v) \in V_{Class} \right\} \cup \left\{ e_{u,v} \mid (\forall u, \forall v) \in V_{Non-class} \right\} \right) \quad (32)$$

$$\max(Denst_{V_{Class}})_{local} \text{ за підходом } \|E(B)\|_0 \rightarrow 1 \text{ зосереджується на граничних}$$

даних. Тобто намагається локально віддалити дані класу від даних поза межами класу. Такий підхід меншою мірою впливає на загальну змінну щільності даних класу $Denst_{V_{Class}}$.

На рівні узагальнення просторове зміщення даних змінює початкову інформативність. Дані внаслідок зміщення в гіперпросторі ознак змінюють як загальну інформативність множини даних, так і безпосередню інформативність кожної одиниці даних.

Зміщення даних в гіперпросторі на глобальному рівні є небажаним, оскільки існує ймовірність зниження інформативної цінності даних.

Таким чином, для розмежування даних в гіперпросторі ознак пріоритет віддається методам локальної зміни компактності даних за напрямком мінімізації двочасткового графа побудованого на МОД загальних нетипових даних $Data_{transition}$:

$$\min \|E(B)\|_0 \quad (33)$$

$$s.t. B \subset MST(Data_{transition}), Data_{transition} = Data_{Class} \cup Data_{Non-class}.$$

За метод оптимізація локальної компактності даних класу в межах граничних екземплярів візьмемо:

$$\min_{\|X\|_0} \|E(B)_{\|X\|_0}\|_0 \quad s.t. M \geq \|X\|_0 \geq k, k \neq 0 \quad (34)$$

Тут $\|\cdot\|_0 \in l_0$ - векторною нормою, яка визначається таким чином $\|E\|_0 = \sum |e|^0$ кількістю ненульових елементів входження у E . E є вектором ребер двочасткового графа.

Зміна простору ознак X здійснюється на γ ознаках. Ці ознаки в загальній постановці є нетиповими та можуть бути викидами, що суттєво впливають на змішування даних класу та даних поза межами класу і утворюють певну множину X_γ . Мінімізація норми ребер двочасткового графа дозволяє детектувати та визначити їхню множину:

$$\left\| E(B)_{\|X\|_0} \right\|_0 \xrightarrow{\|X\|_0 \rightarrow k^+} \min(h), h \geq 1. \quad (35)$$

Детектування наступної ознаки із множини X_γ здійснюється на певним чином зміщених даних.

Для зменшення впливу зміни на загальну інформативність даних, визначаються траєкторії зміни ознак. Введення траєкторій дозволяє встановити усі локальні мінімуми норми $\|E(B)\|_0$. Це в свою чергу дає змогу детектувати множину найбільш впливових ознак на розмежування даних.

Введення траєкторії встановлює послідовність зміни ознак, які в кінцевому випадку приводить до локального мінімуму норми $\min_{\|X\|_0} \|E(B)\|_0$.

В загальному траєкторію зміни ознак представлено у вигляді частково впорядкованої множини $\langle L, \succ \rangle$. Оскільки траєкторія подана у вигляді послідовної зміни, будь-які два елемента впорядковані.

Множина траєкторій зміни ознак має таку форму:

$$\{\langle L, \succ \rangle\} = \left\{ \begin{array}{l} \langle a_1 \succ a_3 \succ a_2 \rangle_1 \\ \langle a_3 \succ a_2 \succ a_1 \succ a_6 \succ a_5 \succ a_4 \rangle_2 \\ \langle a_2 \succ a_1 \succ a_3 \succ a_4 \rangle_3 \\ \langle a_1 \succ a_m \succ \dots \succ a_2 \rangle_4 \\ \vdots \\ \langle a_1 \succ a_2 \succ a_3 \succ \dots \succ a_m \rangle_n \end{array} \right\} \quad (36)$$

Тут за $a_i, i \in [1..m]$, $m \in X$ встановлюється певна унікальна зміна ознаки, яка може бути вбудована у різні траєкторії. В кінцевому випадку $m = r_\alpha$.

Для отримання X_γ , детектуються унікальні елементи траєкторії $\langle L, \succ \rangle$, які є ознаками і відповідно $x \in X_\gamma$.

Метод аналізу складнокласифікованих даних складається з таких кроків:

1. Визначення множини ознак дослідження.
2. Побудова відстаней між даними на основі ознак.
 - 2.1. Побудова мінімального остового дерева.
 - 2.2. Формування множини ребер двочасткового графа.

3. Зміна вибраної ознаки та перехід до кроку 2, перехід до наступного кроку.
4. Визначення впливу зміни ознаки на розмежування даних.
5. Формування кортежу ознак, зміна яких приводить до покращення розмежування.

Встановлені метрики впливу ознак на просторове приграничне розташування складнокласифікованих даних.

Метрика важливості ознаки:

$$\frac{\sum_i 1_{\langle L, > \rangle_i}(\gamma)}{|\{\langle L, > \rangle\}|}, \langle L, > \rangle_i \in \{\langle L, > \rangle\}, i \in [1..|\{\langle L, > \rangle\}|]. \quad (37)$$

Метрика визначає рівень важливості впливу ознаки масиву X_γ та є відношенням кількості траєкторій в яких присутня ознака до загальної кількості траєкторій.

Кардинальне число множини ознак визначається загальною кількістю ознак, що формують множини X_γ , зміна яких призводить до покращення розмежування нетипових даних класів $|X_\gamma|$.

Максимальна довжина траєкторії, яка представлена у вигляді кортежу $|l_{\max}|$.

Максимальна величина редукції міжкласових ребер МОД визначається в абсолютному та відносному представленні:

$$reduc_a = E(B)_0 - E(B)_{\min}, \quad (38)$$

$$reduc_r = \frac{E(B)_0 - E(B)_{\min}}{E(B)_0}. \quad (39)$$

Визначається початкова кількість ребер двочасткового графа, побудованого на вершинах рівних класів $E(B)_0$ та мінімальна кількість ребер при побудові траєкторій редукції $E(B)_{\min}$. Це дозволяє встановити наскільки збільшується компактність даних класу внаслідок зміни ознак у кількості міжкласових ребер.

Відношення кількості унікальних множин даних $\sum \exists! \{x\}$ отриманих з траєкторій $\langle l, > \rangle \rightarrow \{x\}$ до загальної кількості траєкторій:

$$\frac{count_{unic}}{count_{gen}} = \frac{\sum_i 1_{\{\langle L, > \rangle\}}(\{x\}_i)}{|\{\langle L, > \rangle\}|}, \exists! \{x\}_i \leftarrow \{\langle L, > \rangle\}, i \in [1..|\{\langle L, > \rangle\}|]. \quad (40)$$

Співвідношення дозволяє оцінити наскільки широко представлена наявність ознак в траєкторіях загалом.

Використання зміни ознак на множині нетипових даних дозволяє визначити ознаки, які значною мірою впливають на розмежування даних. При цьому дані класу з використанням однокласової класифікації в перехідній зоні утворюють більш компактні області.

Локальне розмежування шляхом мінімізації кількості ребер двочасткового графа визначає викиди ознак за мірою їхнього впливу на розмежування даних класу та даних поза межами класу.

Запропонований підхід дозволяє встановити ознаки серед множин пошуку. Викиди ознак залежать від множини пошуку приналежності даних дослідження відносно класів. Так можна знайти особливі ознаки також серед множини спільних ознак. При цьому отримується гранична межа зменшення двочасткового графу. Вона є спільною незалежно від траєкторії її досягнення та є глобальним мінімумом двочасткового графа на множині ознак дослідження.

Використання пропонованого підходу дозволяє визначити множини ознак, які негативно впливають на розмежування даних і детектувати їх із встановленням міри пливку та отримати їхні розмірні характеристики.

У **шостому розділі** дисертаційної роботи представлена інтелектуальна інформаційна технологія отримання довірчих рішень.

Проекція концептів забезпечення довіри в площину ІС представлена таким чином:

$$\phi_{pr} : (MtMs_o) \mapsto (MtMs_{IIS}) \text{ s.t. } MtMs_{IIS} \subset MtMs_o. \quad (41)$$

Множина складових довіри ІС $MtMs_{IIS}$ визначається як виділення підмножини $MtMs_{IIS} \subset MtMs_o$. Виділення підмножини здійснюється функцією проекції базової множини $MtMs_o$.

Такий стан відповідає довірі до рішень сформованих засобами ІС. ІС в загальному випадку є складовою частиною ІТ, що у втіленні узагальнюючих множин аспектів представлення можна подати у вигляді $IIS \subset IIT$. Відповідно, формування складових довіри у проекції ІС здійснюється з описом підходів, які дозволяють забезпечити практичне втілення та реалізацію цих складових шляхом побудови конкретних механізмів СПР.

В результаті практичної реалізації розробленого метода встановлення відповідності онтології довіри до ІТ та структурованого домену було визначено важливість кожної складової довіри. Важливість складових встановлюється ваговою часткою елементів довіри в загальній сукупності визначеної множини довіри.

Відповідно, на прикладному рівні методів СПР складові елементи довіри визначається таким чином:

– Інтерпретованість, пояснювальність, транспарентність, простежуваність. Базовим узагальнюючим концептом є транспарентність. Транспарентність визначається стосовно стейкхолдера як здатність системи бути зрозумілою та пояснювальною за результатами прийнятих рішень і за кроками прийняття рішень. Відповідно, метод ІС з прийняття рішення розбивається на послідовні кроки виконання з вимогою зрозумілості та прозорості їх виконання відносно стейкхолдера. Визначається інтерпретованість як внутрішній стан системи, що відповідає здатності бути зрозумілою стейкхолдеру. Пояснювальність як зовнішній стан системи, що відповідає здатності бути зрозумілою стейкхолдеру. Прийняття рішень в ІС здійснюється СПР, яка відповідним чином будується та валідується. Інтерпретованість СПР описується рівнем інтерпретованості алгоритму, який використовується для її побудови. На рівні розроблених методів складові реалізуються шляхом застосування групового прийняття рішення у вигляді ансамблю СПР. Консолідовані рішення ансамблю інтерпретуються за СПР, яка володіє максимальною інтерпретованістю стосовно СПР ансамблю.

– Всебічність розгляду. Множина даних за консолідованими рішеннями є множиною довірчих рішень на основі всебічності розгляду прийняття рішень. Також формується множина даних, відносно якої наявні СПР, які мають альтернативні рішення по відношенню до рішень СПР групи. Така множина даних не може вважатися довірчою у відношенні до рішень групи СПР. Таким чином, формується загальне рішення ансамблю СПР в сегменті даних альтернативних рішень.

– Аналізованість даних. В межах дослідження аналізованість даних визначається встановленням рівня впливу ознак даних на розмежування складнокласифікованих даних, які формуються множиною нетипових даних.

Множина принципів сформована використовуючи людиноцентровану концепцію, яка є визначальною концепцією соціотехнологічного розвитку суспільства. Далі використовуючи розроблену онтологію довіри були встановлені складові довіри.

Використовуючи метод визначення відповідності складових онтології довіри до інших узагальнень етичних принципів довіри було встановлено множину етичних принципів, які лягти в основу проєктування та розробки інтелектуальної інформаційної технології в частині прийняття довірчих рішень. Використовуючи синтаксичний, структурний та семантичний рівні встановлення відповідності етичних принципів складовим структурованого домену було визначено важливість кожного принципу відносно загальної множини принципів.

Забезпечення отримання довірчих рішень здійснюється шляхом побудови інформаційної технології на визначених етичних принципах. Структура формування інформаційної технології представлена на рис. 3.

Розроблена інформаційна технологія забезпечення довіри до інтелектуальних інформаційних систем у вигляді взаємодії запропонованих методів представлено на рис. 4.

Інтелектуальна інформаційна технологія генерації довірчих рішень описується набором функцій:

– формування сегменту інформації на основі вхідної інформації рішення відносно якої є довірчими та встановлені складовими інтелектуальної системи прийняття рішень на основі визначених етичних принципів;

– формування сегменту інформації на основі вхідної інформації рішення відносно якої є недовірчими, що встановлено складовими інтелектуальної системи прийняття рішень, така інформація визначається як складнокласифікована;

– отримання висновку щодо якісних метрик класифікації вхідної інформації на загальній множині інформації включаючи довірчу інформацію та складнокласифіковану інформацію;

– отримання висновку щодо ознак, які мають визначальний вплив на розмежування складнокласифікованої інформації залежно від класів приналежності інформації;

– отримання висновку щодо просторового розташування класів інформації в гіперпросторі ознак та візуальному представленні для візуального аналізу та контролю.

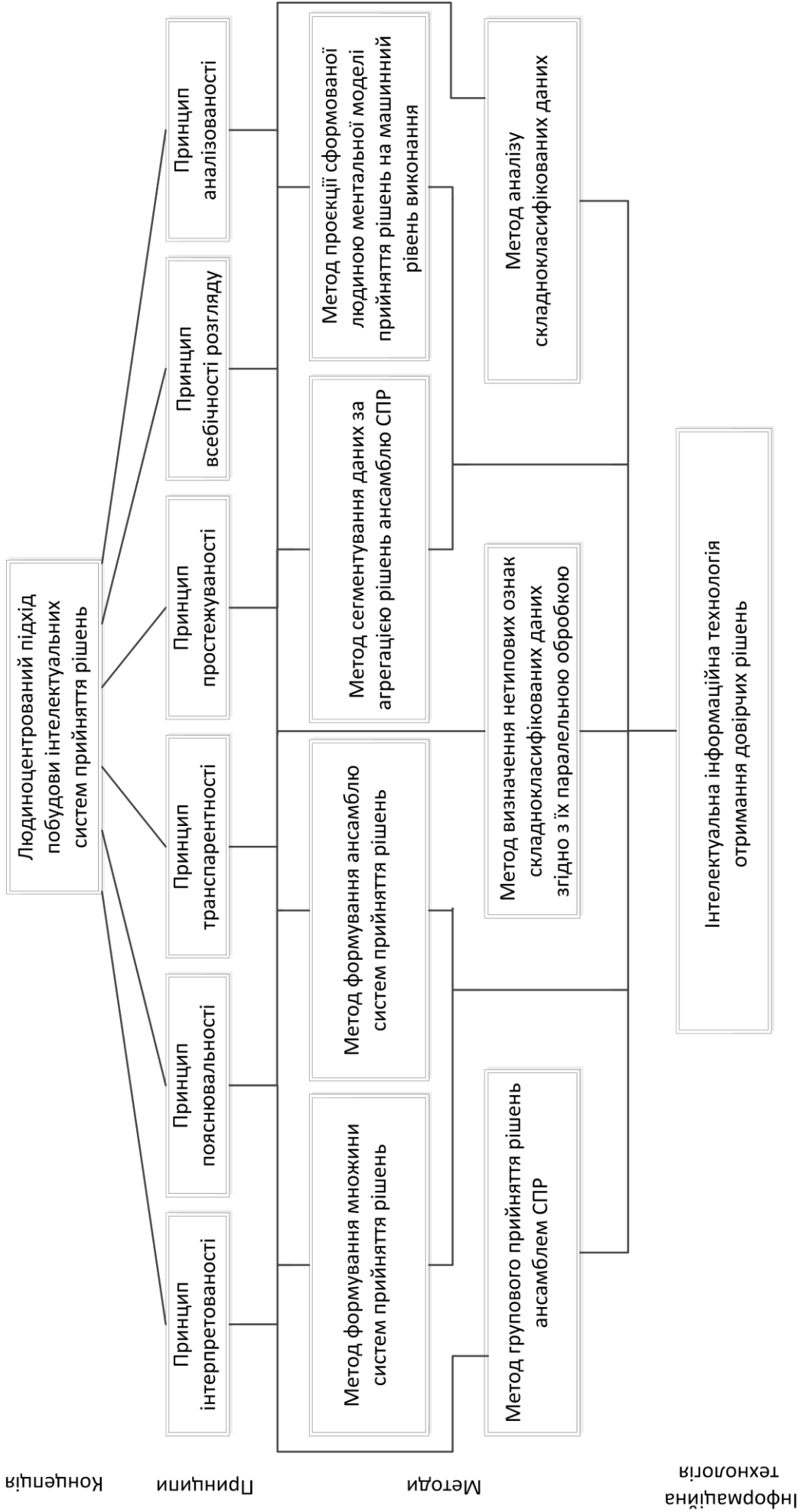


Рис. 3. Структура формування інтелектуальної інформаційної технології отримання довірчих рішень

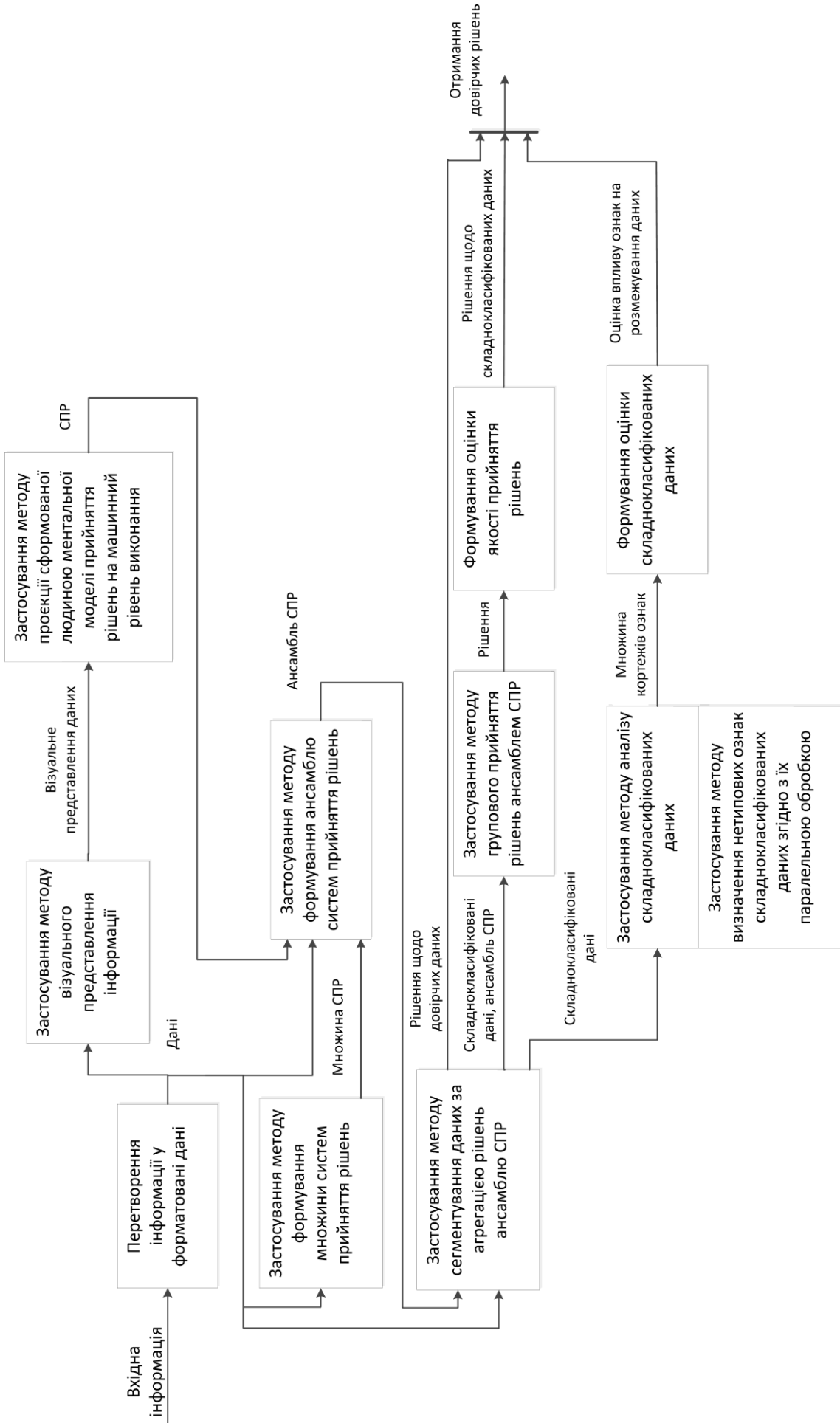


Рис. 4. Структура взаємодії складових інтелектуальної інформаційної технології отримання довірчих рішень

З метою забезпечення реалізації та практичного втілення встановленої множини етичних принципів розроблено низку методів практичного спрямування та функціонального призначення. Інтелектуальна інформаційна технологія представляє собою цілісну форму об'єднання розроблених методів в єдину систему, яка підпорядкована забезпеченню довіри до рішень отриманих за допомогою інтелектуальної інформаційної системи.

Отримання довірчих рішень визначає загальну довіру до інтелектуальної інформаційної технології в межах встановлених етичних принципів за людиноцентрованим підходом.

Розроблено метод групового прийняття рішень ансамблем систем прийняття рішень. Сегментування даних за ознаками довіри здійснюється шляхом використання групування СПР та формування ансамблів. Ініціалізується пул претендентів класифікаторів $C_0 = \{c\}_{i=1}^n, n \in \mathbb{N}$, який складається з класифікаторів апріорно потенційних претендентів для включення до ансамблю. Із використанням цих алгоритмів класифікаторів будуються СПР даних для тренування $Data_{train}$.

Метод групового прийняття рішень ансамблем СПР складається з таких кроків:

- застосування класифікаторів ансамблю до вхідних даних;
- формування множини комбінації поєднань рішень класифікаторів в межах ансамблю;
- застосування агрегативного методу до поєднань рішень множин комбінацій класифікаторів;
- визначення найкращої комбінації класифікаторів за якісним показником та необхідним рівнем інтерпретованості.

Валідація даних для побудови СПР класифікатора здійснюється шляхом використання побудованої СПР на всій множині тренувальних даних $Data_{train} \equiv Data_{test}$. В загальному випадку процес формування даних тренування представляється таким чином:

$$\begin{aligned} f_n(Data_{train\ n}) &= f[f_{n-1}(Data_{train\ n-1})] \rightarrow \min(Err(m_c(Data_{train\ n}))) \\ s.t. \min(Err(m_c(Data_{train\ n}))) &\rightarrow 0 \end{aligned} \quad (42)$$

З базової множини даних для тренування утворюються множини валідованих даних для побудови СПР кожного з класифікаторів $Data_{train,c} \subset Data_{train}, c \in C_0$.

Наступним кроком є формування ансамблю з пулу потенційних класифікаторів C_0 . В основі відбору лежить здатність кожного з класифікаторів привнести в ансамбль свою унікальну здатність до розпізнання інформативності, яку містять дані. Таким чином, в ансамблі СПР кожний з класифікаторів доповнює один одного з метою виявлення помилок класифікації, всебічності розгляду даних, підсилення слабких сторін інших класифікаторів та інше. Результатом такого відбору класифікаторів є сформований ансамбль СПР $M = \{m\}_{i=1}^n, n = |C|$ класифікаторів $C \subseteq C_0$. В загальному випадку множина відібраних класифікаторів є підмножиною пулу потенційних класифікаторів або відповідає їй.

На основі ансамблю СПР M формується множина можливих комбінацій рішень СПР $CMB = \{cmb\}_{i=1}^n$, $n = |P^+(M)|$. Ця множина є вихідною для формування комбінації поєднань групувань рішень за їхньою подібністю. Кожний елемент комбінаторних поєднань множини CMB рішень СПР M подається на агрегацію за групуванням рішень. Рішення СПР, які входять до знайдених груп об'єднуються в рішення цих груп. Так формують рішення наступного рівня ієрархії шляхом делегування рішень окремих СПР, як рішення групи в яку вони входять. За алгоритмом k-means знаходиться центр кожної групи, який представляє собою узагальнене рішення СПР цієї групи (рис. 5.). Використовується трьохшарова структура ієрархії формування рішень ансамблю з двома міжшаровими переходами. Загальна множина рішень за ієрархічною агрегацією $AGG = \{agg\}_{i=1}^n$, $n = |P^+(M)|$ буде відповідати множині комбінаторних груп $|AGG| = |CMB|$.

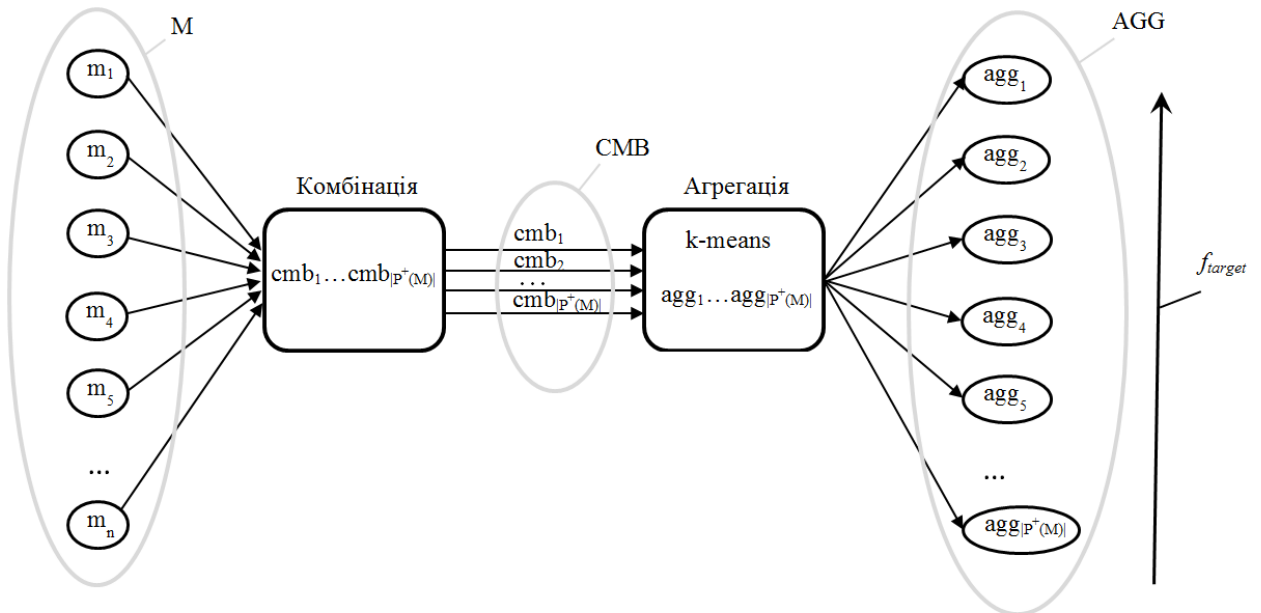


Рис. 5. Схематичне представлення методу отримання рішень з вибором оптимального набору з ансамблю СПР за цільовим показником якості

Загальна множина тестових даних сегментується на типові дані $Data_{test_valid}$ консолідованих рішень ансамблю та нетипові дані D_{Trans} наявності альтернативних рішень ансамблю СПР, $Data_{Test_valid} \cup Data_{Trans} = Data_{Test}$. Рішення на множині тестових даних розділюється на рішення відповідних підмножин $AGG|_{Data_{Test_valid}} \cup AGG|_{Data_{Trans}} = AGG|_{Data_{Test}}$. Агреговані рішення типових даних є консолідовані та за рішенням ансамблю – довірчими. Інша множина нетипових даних не є довірчою.

З ансамблю СПР утворюється множина дискретних рішень з агрегативним ієрархічним підходом на множині нетипових навчальних даних. Множина дискретних рішень відповідає множині можливих комбінацій рішень СПР, які утворюють ансамбль. Із множини загальних рішень вибирається така комбінація СПР групування рішень, яка максимізує певний цільовий показник якості

класифікації. Таким чином, результуючий показник якості класифікації усієї множини даних тестування є поєднанням результатів тестування на множині консолідованих рішень та на множині нетипових даних тобто наявності альтернативних рішень:

$$f_{target}(Data_{Test}) = f\left(f_{target}(Data_{Test_valid}), \max(f_{target}(Data_{Trans}))\right). \quad (43)$$

За кінцевим результатом тестова множина даних розділюється на сегменти даних, формується сегмент довірчих рішень та сегмент даних рішенням щодо яких не можна довіряти. Серед множини нетипових даних максимізується цільовий показник якості класифікації шляхом вибору рішень СПР з множини ансамблю. Цільовий параметр якості буде визначатись рішеннями усіх СПР з ансамблю на множині довірчих даних та відібраними рішеннями на множині нетипових (недовірчих) даних:

$$dom f_{target} = f\left(\bigcup_{i=1}^{|C|} dom f_{target} \Big|_{m, \forall m \in M}, \bigcup_{i=1}^n dom f_{target} \Big|_{m, \forall m \in M, n \in P^+(M)}\right). \quad (44)$$

Розроблено метод визначення нетипових ознак складнокласифікованих даних згідно з їх паралельною обробкою. Мінімізація двочасткового графа нетипових даних $\min_{\|X\|_0} \|E(B)_{\|X\|_0}\|_0$ s.t. $M \geq \|X\|_0 \geq k, k \neq 0$ здійснюється з метою детектування γ

ознак, тобто ознак, наявність яких негативно впливає на групування даних класу. Мінімізація здійснюється шляхом ущільнення даних класу в зоні нетипових даних $\max(Denst_{V_{Class}})_{local}$ та отримання множини $X_\gamma = \{\gamma_i\}_{i=1}^{\gamma}$. Оскільки в загальному випадку існує множина локальних мінімумів при мінімізації множини ребер двочасткового графа $(B)_{\|X\|_0}$ шляхів *direct* з досягнення кожного окремого мінімуму

$Local = \{local\}_{i=1}^n, n \in \mathbb{N}$ може існувати декілька $Direct_{\min\|E(B)\|} = \{direct\}_{i=1}^n, n \in \mathbb{N}$. В

загальному випадку записується у такому вигляді

$Direct \times Local$ s.t. $Local = \min\|E(B)_{X_\gamma}\|$. Певна множина шляхів досягнення

локального мінімуму формується шляхом детектування γ ознак.

Зміна співвідношень взаємного положення даних в гіперплощині ознак при зміні ознак $a_i, i \in [1..m], m \in X$ визначається окремим станом при кожному шляху

direct досягнення мінімуму $\min\|E(B)\|$. Відповідно, шлях *direct* характеризується

множиною змін ознак $direct \rightarrow f\{a\}$. Проте на множині *Direct* при досягненні

локального мінімуму $direct_i \Big|_{\min\|E(B)\|} \neq direct_j \Big|_{\min\|E(B)\|}, i \neq j \in |Direct|$ при збігові множин

змін ознак $\{a\}_{direct_i} \equiv \{a\}_{direct_j}$ за тих же умов. Зміна ознак $\{a\}$ призводить до

дискретизації станів простору даних $(V \subset R^{CS}, dist)$ за характерним взаємним розташуванням даних. Відповідно, множину шляхів представлено у вигляді

множини векторів, а саме послідовних кортежів зміни ознак $L \equiv \overline{Direct}$. В загальному випадку траєкторії змін ознак було представлено у вигляді частково впорядкованої множини $\langle L, \succ \rangle$. Кожна траєкторія є унікальним вектором лінійно впорядкованим, тому в межах узагальнення траєкторії змін ознак набуде вигляду строго впорядкованої множини $\langle L, > \rangle$.

Для отримання траєкторій $\langle L, > \rangle$ з пошуку множини мінімумів $\min \|E(B)\|$ необхідно виконати наступні кроки методу формування множини γ ознак $X_\gamma \subset X$.

Метод визначення нетипових ознак складнокласифікованих даних згідно з їх паралельною обробкою складається з таких кроків:

- визначення множини ознак, стосовно яких здійснюється дослідження рівня впливу на розмежування даних;
- визначення множини даних відносно класу дослідження;
- застосування послідовної структури перетворення даних та ознак області дослідження;
- розпаралелювання множини ознак на структури паралельного виконання;
- формування множини траєкторій зміни ознак з мінімізацією множини ребер двочасткового графа.

Опишемо процес виконання кожного кроку детально. Сформуємо початкову множину ознак, яка в загальному випадку $XR_0 = X$. Необхідність формування початкової множини ознак зумовлена проєкцією на множину дослідження.

Залежно від проєкції в межах однокласової класифікації вибирається один з варіантів:

1. Область дослідження визначається усією множиною ознак X нетипових даних D_{Trans} , $XR_0 = X$. $X = X|_{D_{Trans}}$, оскільки множина ознак описується в межах множини нетипових даних.

2. Множина ознак формується на основі даних, які є підмножиною нетипових даних D_{Trans} і належать до даних класу D_{Class} , $D_{Class} \subset D_{Trans}$. В такому випадку $XR_0 = X|_{D_{Class}}$.

3. Множина ознак формується на основі даних, які є підмножиною нетипових даних D_{Trans} та належать до даних поза межами класу $D_{Non-class}$, $D_{Non-class} \subset D_{Trans}$. В такому випадку $XR_0 = X|_{D_{Non-class}}$.

Крім формування початкової множини ознак дослідження на основі даних за умовою відношення даних до класу, початкову множину можна сформувати за відношенням до даних.

1. Множина ознак формується на основі ознак, які належать тільки нетиповим даним $XR_0 = \sigma_{x:\forall x \notin X|_D} (X|_{D_{Trans}})$, $D_{Trans} \subset D$, $X|_{D_{Trans}} \setminus X|_D \neq \emptyset$.

2. Множина ознак формується на основі ознак, які належать тільки нетиповим даним та відносяться до даних класу D_{Class} $XR_0 = \sigma_{x:\forall x \notin X|_D \wedge \forall x \in X|_{D_{Class}}} (X|_{D_{Trans}})$, $D_{Trans} \subset D$, $X|_{D_{Trans}} \setminus X|_D \neq \emptyset$.

3. Множина ознак формується на основі ознак, які належать тільки нетиповим даним та відносяться до даних поза межами класу $D_{Non-class}$

$$XR_0 = \sigma_{x:\forall x \notin X|_D \wedge \forall x \in X|_{D_{Non-class}}} \left(X|_{D_{Trans}} \right), \quad D_{Trans} \subset D, \quad X|_{D_{Trans}} \setminus X|_D \notin \emptyset.$$

Подальшим розвитком функції відбору множини ознак дослідження є формування початкової множини на основі відношення до даних.

1. Множина ознак формується на основі ознак, які належать тільки нетиповим даним та відносяться до даних класу D_{Class} та не відносяться до даних поза межами класу

$$D_{NonClass} \quad XR_0 = \sigma_{x:\forall x \notin X|_D \wedge \forall x \in X|_{D_{Class}} \wedge \forall x \notin X|_{D_{Non-class}}} \left(X|_{D_{Trans}} \right), \quad D_{Trans} \subset D,$$

$$X|_{D_{Trans}} \setminus X|_D \notin \emptyset.$$

2. Множина ознак формується на основі ознак, які належать тільки нетиповим даним та відносяться до даних поза межами класу $D_{Non-class}$ та не відносяться до даних класу

$$D_{Class} \quad XR_0 = \sigma_{x:\forall x \notin X|_D \wedge \forall x \notin X|_{D_{Class}} \wedge \forall x \in X|_{D_{Non-class}}} \left(X|_{D_{Trans}} \right), \quad D_{Trans} \subset D,$$

$$X|_{D_{Trans}} \setminus X|_D \notin \emptyset.$$

3. Множина ознак формується на основі ознак, які належать тільки нетиповим даним та відносяться до даних поза межами класу $D_{Non-class}$ і відносяться до даних класу D_{Class} , тобто виключно на спільних ознаках

$$XR_0 = \sigma_{x:\forall x \notin X|_D \wedge \forall x \in X|_{D_{Class}} \wedge \forall x \in X|_{D_{Non-class}}} \left(X|_{D_{Trans}} \right), \quad D_{Trans} \subset D, \quad X|_{D_{Trans}} \setminus X|_D \notin \emptyset.$$

Наступний крок складається з таких етапів:

– на основі вибраної початкової множини ознак, що визначається областю дослідження, обчислюється матриця евклідових відстаней $Dist_0$ нетипових даних D_{Trans} .

– побудова мінімального остового дерева MST_0 на обчислені матриці відстаней $Dist_0$.

– виділення множини ребер $E(B)$ двочасткового графа B_0 з вершинами підмножини вершин мінімального остового дерева $V_B \subset V_{MST}$, де ребра двочасткового графа визначаються інцидентними вершинами, які належать до множини даних класу та множини даних поза межами класу $\{e_{u,v} | \forall u \in V_{Class}, \forall v \in V_{Non-class}\}$.

Перетворення вхідних даних, які визначають область дослідження ознак виділяє необхідну інформативність даних, що проявляється у побудові двочасткового графа в межах концепції однокласової класифікації.

Необхідним є виділення ознаки дослідження x з ініціалізованої множини ознак XR_0 та формування нового простору ознак $XR_{i,j} = XR_{i-1} \setminus \{x_j\}$, $i = 0..|l|$, $\forall l \in L$, $j = 1..|XR_{i-1}|$. На цьому етапі кількість досліджуваних ознак визначається множиною XR_{i-1} , тобто множиною ознак попереднього етапу дослідження формування ознак.

Кожна ознака досліджується окремим етапом незалежно від інших ознак, що може виконуватись у паралельному режимі. Кількість паралельних потоків $|XR_{i-1}| - 1$.

Обрахунок блока послідовної структури, який в загальному представлений таким чином: $Dist(XR) \rightarrow MST \rightarrow B \rightarrow E(B)$.

Рішення TF про належність ознаки $x_{i,j}$ до множини γ ознак визначається співвідношенням:

$$TF : |E(B_{i-1})| - |E(B_i)| > 0. \quad (45)$$

Далі застосовуємо формування кортежу ознак послідовного збільшення компактності класу $\langle TR_{i,j} \rangle = \langle TR_{i-1,j} \rangle \cup x_j$. Умовою завершення формування кортежу ознак є умова обернена до рішення TF , тобто до наявності ознаки формування кортежу:

$$|E(B_{i-1})| - |E(B_i)| \leq 0. \quad (46)$$

Сформований кортеж відповідає траєкторії $l_k \in L, k = 1..|L|$, $l_k = \langle TR_{i,j} \rangle$ пошуку локального мінімуму $\min \|E(B)\|$ наявності міжкласових зв'язків $link = \{Class, Non - class\}$ на основі МОД.

Об'єднуємо траєкторії за елементами кортежів у множину X_γ за унікальними ознаками:

$$X_\gamma = \sigma_{\exists!x, \forall x \in XR_0} \left(\bigcup \langle TR \rangle \right). \quad (47)$$

Таким чином, формується множина ознак X_γ , тобто ознак, зміна яких призводить до просторового ущільнення нетипових даних шляхом зменшення міжкласових ребер мінімального остового дерева побудованого на цих даних.

У **сьомому розділі** дисертаційної роботи наведені експериментальні дослідження методів сегментації даних за складовими довіри та аналізу ознак нетипових даних.

Виділення множини даних довірчих рішень із загальної множини даних супроводжується розділенням даних на два сегменти, множини даних довірчих рішень та множини нетипових даних рішення щодо яких є недовірчими. Проте якісні показники класифікації даних застосовуються до загальної множини даних, об'єднуючи значення показники для сегментів даних. Для цього використовуються представлені в попередніх розділах модифіковані показники точності та повноти.

Довіра відносно рішень щодо даних формується на основі консолідованості рішень СПР. Це забезпечується не тільки архітектурними та проєктними підходами, а також показниками якості. Зі збільшенням кількості класифікаторів, які представлені СПР в ансамблі множина довірчих даних зменшується з відповідним збільшенням довіри. Наприклад, 2 СПР в ансамблі дають консолідовані рішення 97.9% довірчих даних, 3 СПР відповідно 96.8%, 6 СПР відповідно 93.6%. В мінімальному значенні величина довірчих даних досягала 89.3%. В даному випадку розглядається множина консолідованих рішень. Поряд з тим консолідованість рішень не виключає також наявність помилок у множині довірчих рішень. Таким

чином, збільшення розміру ансамблю СПР підвищує довіру до рішень, а також впливає на об'єктивність та всебічність розгляду рішення щодо елемента даних.

Рішення СПР дають ймовірнісну оцінку належності елемента даних до класу. Розглядається однокласова класифікація, де рішення представлені ймовірнісною оцінкою в межах $[1, -1]$. Максимальне значення оцінки рівне одиниці у випадку рішення приналежності до класу та «мінус» одиниці – у випадку рішення не належності до класу. За такими ж оцінками визначають групування рішень та результуюче рішення поєднання рішень СПР щодо кожної одиниці даних. Задавши вектор оцінки приналежності до класу кожної СПР в поєднанні та вектор оригінальних значень відношення до класу на множині нетипових даних, отримано значення корельованості рішення СПР. Для визначення ефективності метода ієрархічної групуваності прийняття рішень проведені експериментальні дослідження для дев'яти класів та представленні зведені результати за макрохарактеристиками для десяти кращих поєднань (Таблиця 1).

Таблиця 1.

Результати експериментальних досліджень з однокласової класифікації за макропоказниками поєднань рішень СПР

№	Отриманні значення для класів, точність /повнота								
	1	2	3	4	5	6	7	8	9
1	0.9297/ 0.9297	0.9665/ 0.7089	0.9705/ 0.7531	0.9949/ 0.9222	0.9778/ 0.713	0.9685/ 0.6615	0.8541/ 0.5447	0.9972/ 0.8889	0.9979/ 0.9
2	0.9062/ 0.9062	0.9601/ 0.8212	0.9697/ 0.8264	0.9934/ 0.9	0.9764/ 0.6944	0.9672/ 0.6462	0.8375/ 0.5743	0.9972/ 0.8889	0.9972/ 0.8667
3	0.8986/ 0.951	0.9592/ 0.6392	0.9685/ 0.7346	0.992/ 0.8778	0.971/ 0.6204	0.9652/ 0.6231	0.8375/ 0.5743	0.9972/ 0.8889	0.9965/ 0.8333
4	0.8986/ 0.951	0.9591/ 0.7524	0.9678/ 0.7284	0.9906/ 0.8556	0.9569/ 0.7122	0.9645/ 0.6154	0.8375/ 0.5743	0.9965/ 0.8611	0.9958/ 0.8
5	0.8937/ 0.9282	0.9572/ 0.6203	0.9644/ 0.6975	0.9898/ 0.8444	0.9448/ 0.8126	0.9632/ 0.6	0.8317/ 0.6039	0.9965/ 0.8611	0.9958/ 0.8
6	0.8937/ 0.9282	0.9572/ 0.6203	0.9597/ 0.6543	0.9877/ 0.8111	0.943/ 0.8033	0.9626/ 0.5923	0.817/ 0.6479	0.9958/ 0.8333	0.9958/ 0.8
7	0.8778/ 0.9275	0.9571/ 0.7397	0.9566/ 0.74	0.9849/ 0.7667	0.943/ 0.8033	0.9619/ 0.5846	0.8157/ 0.6183	0.9951/ 0.8056	0.9944/ 0.7333
8	0.8743/ 0.9723	0.9566/ 0.6139	0.9546/ 0.7276	0.9842/ 0.7556	0.9411/ 0.794	0.9606/ 0.5692	0.8157/ 0.6183	0.9944/ 0.7778	0.993/ 0.6667
9	0.8689/ 0.9495	0.9561/ 0.7334	0.9484/ 0.7577	0.9831/ 0.9326	0.9391/ 0.7848	0.9547/ 0.7069	0.8143/ 0.5887	0.9944/ 0.7778	0.9923/ 0.6333
10	0.8689/ 0.9495	0.9561/ 0.7334	0.948/ 0.6906	0.982/ 0.9215	0.9371/ 0.7755	0.944/ 0.6531	0.7962/ 0.6471	0.9944/ 0.7778	0.9923/ 0.6333

В результаті проведених експериментів відзначається загальне покращення поєднань показників точності та повноти відносно базових рішень класифікаторів.

При високих показниках базових СПР покращення знаходиться для точності в межах 1- 3.5 %, повноти 3-9 %. При визначенні поєднання з одночасним визначенням характеристик точність може знижуватись до 1 % при підвищенні повноти до 7 %.

Результати експериментальних досліджень показали, що розроблений метод сегментування даних за агрегацією результатів роботи ансамблю систем прийняття рішень дозволив сформувати сегменти даних на основі довіри до рішень ансамблю СПР. Запропонований метод дозволив практично реалізувати механізми, які імплементують визначені етичні принципи, що забезпечують формування концепту довіри до рішень ансамблю СПР.

Із загальної множини даних дослідження виділено множину, щодо якої встановлено реалізацію за сукупністю визначених етичних принципів концепту довіри. Відповідно, на основі довіри щодо рішень СПР класифікаторів такі рішення визначено як довірчі. Мінімальне значення величини довірчих даних становить 89,3 %.

Реалізація запропонованого метода дозволила сформувати множину даних, які характеризуються наявністю альтернативних рішень в ансамблі СПР. Такі дані встановлені як недовірчі за невизначеністю реалізації складових концепту довіри. Ці дані характеризуються як нетипові по відношенню до довірчих даних за сукупністю ознак.

Проведення експериментальних досліджень методу групового прийняття рішень ансамблем систем прийняття рішень на множині нетипових складнокласифікованих даних та інтелектуальної інформаційної технології отримання довірчих рішень підтверджують ефективність СПР та запропонованих методів.

Результати досліджень показали загальну ефективність порівняно з базовими СПР ансамблю та показниками точності в межах 3.5 %, повноти в межах 9 % на загальній множині даних, що дозволяє використовувати запропоновану технологію для задач класифікації на базі СПР. Метод аналізу ознак складнокласифікованих даних дозволив встановити множину ознак, які мають визначальний вплив на групованість даних. Міру впливу ознаки визначають за сукупністю метрик. Множина ознак, які значною мірою впливають на розмежування даних складає менше ніж 1 % від загальної множини ознак нетипових даних.

У **додатках** наведено результати експериментальних досліджень. Також наведено впровадження результатів дисертаційних досліджень.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальну науково-прикладну проблему підвищення рівня довіри до інтелектуальних інформаційних технологій в частині прийняття ними рішень на практичному рівні шляхом розроблення теоретичних та прикладних засад інтелектуальної інформаційної технології отримання довірчих рішень за людиноцентрованим підходом, яка забезпечує довіру до рішень,

отриманих інтелектуальними інформаційними системами відповідно до сукупності етичних принципів.

Згідно із поставленою метою, в дисертаційній роботі вирішено наступні завдання та отримано відповідні наукові та практичні результати:

1. Проведено дослідження стану сучасного розвитку інформаційних технологій із використанням інтелектуальних складових. Проаналізовано вплив, застосування та інтегрування інтелектуальних інформаційних технологій у сфери життя як окремої людини, так і суспільства загалом. Встановлено, що інтелектуальні інформаційні технології проявляються не тільки у вигляді технологічних рішень, а й також є визначальним фактором розвитку та трансформації суспільства. Інтеграцію інтелектуальних інформаційних систем слід розглядати з позиції соціотехнологічних систем. Встановлено, що основою широкого використання інтелектуальних інформаційних технологій та делегування можливості прийняття рішень є довіра до усіх проявів та складових інтелектуальних інформаційних технологій. Визначено, що необхідним та актуальним напрямом є дослідження методів формування та забезпечення практичної реалізації концепту довіри до інтелектуальних інформаційних технологій, який впливає із дотримання етичних принципів, визначених на основі людиноцентрованого підходу.

2. Запропоновано онтологію довіри до інтелектуальних інформаційних технологій. На основі онтології сформовано множину вимог до програмного рівня інтелектуальних інформаційних технологій, а саме методів формування систем прийняття рішень з генерації довірчих рішень. На основі розробленої структури здійснено декомпозицію концепту довіри. Набув подальшого розвитку метод визначення відповідності складових онтології довіри до інших узагальнень етичних принципів довіри. В основі методу лежить об'єднання трьох незалежних рівнів (синтаксичний, структурний, семантичний) встановлення відповідності елементів множини декомпозиції та складових структурованого домену, який в межах роботи представлено у вигляді структурованого корпусу узагальнених документів світового дискурсу людиноцентрованого підходу етичних інтелектуальних інформаційних технологій за фактором довіри.

3. *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 забезпечення належних показників якості класифікації.

4. *Вперше розроблено* метод проєкції сформованої людиною ментальної моделі прийняття рішень на машинний рівень виконання. Метод був розроблений з метою забезпечення складових довіри, які визначаються стейкхолдерами, тобто зацікавленою стороною (людиною) і дозволяє ментальне представлення прийняття рішення сформоване людиною з використанням візуальної аналітики спроектувати на рівень формального представлення, тобто безпосереднє використання здійснюється обчислювальною системою на машинному рівні. Таким чином, рівень

довіри до системи прийняття рішень підвищується значною мірою, оскільки представлення (ментальне) створюється безпосередньо самою людиною і в подальшому проєктується на машинний рівень з формуванням представника – формального представлення. Метод реалізований в межах запропонованої концепції розділення сфери створення представлення (аналітичні можливості людини) та сфери практичного використання (обчислювальні можливості машини).

5. *Вперше розроблено* 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.

6. *Набув подальшого розвитку* метод визначення нетипових ознак складнокласифікованих даних згідно з їх паралельною обробкою. Нетипові дані є складнокласифікованими за підходами класифікації. Однак, аналіз цих даних є важливим з точки зору детектування даних для досліджень зі змінними значеннями ознак. Метод дозволяє визначити множину ознак, які значною мірою впливають на розмежування нетипових даних.

7. *Набув подальшого розвитку* 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 із складовими ансамблями. Ієрархічна структура дозволяє враховувати рішення кожної з системи прийняття рішень множини ансамблів та максимізувати рішення класифікаторів за показниками якості.

8. *Удосконалено* метод формування ансамблів систем прийняття рішень, який відрізняється від відомих тим, що складові ансамблі формуються залежно від необхідного рівня інтерпретованості рішень ансамблів, завдяки чому з'являється можливість отримати рішення ансамблів з належним рівнем довіри. Метод дозволяє ефективно використати як здатність систем прийняття рішень до узагальнення, так і здатність індивідуальної системи прийняття рішень враховувати інформативність окремих складнокласифікованих даних, які визначають необхідність залучення кожної системи прийняття рішень в ансамбль як форми прийняття групового рішення.

9. *Удосконалено* метод сегментування даних за агрегацією результатів роботи ансамблів систем прийняття рішень, перевага якого над відомими методами полягає в тому, що розв'язки складових ансамблів визначено критерієм сегментування даних, стосовно яких приймаються рішення, що уможливило формування сегменту даних, рішення стосовно яких відповідають складовим довіри і є одноголосними, та сегмент складнокласифікованих даних, щодо яких присутня невизначеність при прийнятті рішень. Метод дозволив виділити два сегменти даних: сегмент типових даних, рішення щодо яких характеризуються складовими довіри, та сегмент нетипових даних, щодо яких не можна повною мірою застосовувати та практично реалізовувати виконання складових довіри.

10. Запропонована інтелектуальна інформаційна технологія отримання довірчих рішень за показниками довіри та якісними показниками набула *практичної*

реалізації, що дозволяє отримувати множини довірчих рішень, застосовуючи ансамблевий підхід сформувати отримувати необхідне дискримінантне розмежування рішень за цільовими показниками якості. Отримані набори ознак для слабороздільних даних, які визначають максимально можливу роздільність, та дозволяють встановити міру впливу ознак на розмежуваність даних для ознак певної цільової групи. Результати проведених практичних досліджень дозволили визначити складові концепту довіри з проєкцією на програмний рівень побудови систем прийняття рішень. Практичне застосування інтелектуальної інформаційної технології отримання довірчих рішень дозволило виділити сегмент даних, рішення щодо яких характеризуються високим ступенем довіри, та сегмент нетипових даних поблизу зони розмежування. Поряд з тим вдалося отримати високі показники якості розмежування складнороздільних даних з виділенням ознак із суттєвим значним впливом на розмежування даних.

11. Проведені дослідження підтвердили *практичну цінність* розробленої інтелектуальної інформаційної технології отримання довірчих рішень, яка дала можливість забезпечити довіру до інтелектуальних інформаційних систем в частині прийнятих рішень відповідно до етичних принципів за людиноцентрованим підходом. Результати експериментальних досліджень показали, що інтелектуальна інформаційна технологія забезпечує формування сегментів інформації на основі складових довіри. Застосування розробленої інтелектуальної інформаційної дозволяє отримати сегмент довірчих даних в межах 76–86 %, що дало можливість із загальної множини вхідної інформації дослідження виділити множини інформації, щодо якої встановлено довіру за сукупністю визначених етичних принципів. Використання методу групового прийняття рішень в інтелектуальній інформаційній системі дозволило підвищити точність класифікації в максимальних значеннях до 3 %, повноту – до 9 %, що підтверджує ефективність класифікації за якісними показниками та його необхідність застосування в інтелектуальній інформаційній технології як складової частини з формування довірчих рішень. В сегменті складнокласифікованої інформації інтелектуальна інформаційна технологія дозволяє визначити множини ознак, які мають визначальний вплив на групуваність інформації. Метод аналізу складнокласифікованих даних дозволяє визначити найбільш впливові ознаки на розмежування інформації в сегменті складнокласифікованих даних, множина яких становить близько 4.2 %. Експериментальні дослідження підтвердили ефективність застосування інтелектуальної інформаційної технології отримання довірчих рішень, що дозволяє сформувати множини інформації, рішення щодо якої визначаються як довірчими, а також отримати високі якісні показники класифікації на загальній множині вхідної інформації, та провести аналіз ознак в сегменті складнокласифікованої інформації.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті у наукових виданнях, включених до Переліку наукових фахових видань України:

1. Бармак О. В., Калита О. Д., Манзюк Е. А. Аналіз моделей для розпізнавання мімічних проявів емоцій. *Вимірювальна та обчислювальна техніка в технологічних процесах*. 2020. №1. С. 77–83. URL: <https://doi.org/10.31891/2219-9365-2020-65-1-12>.
2. Бармак О. В., Крак Ю. В., Манзюк Е. А. Характеристика для вибору моделей у ансамблі класифікаторів. *Проблеми програмування*. 2018. №2–3. С. 171–179. URL: <https://doi.org/10.15407/pp2018.02.171>.
3. Бармак О. В., Манзюк Е. А., Калита О. Д., Крак Ю. В., Кузнецов В. О., Куляс А. І. Розпізнавання емоційних проявів за групуванням скупченостей характерних мімічних станів обличчя людини. *Проблеми програмування*. 2020. №2–3. С. 173–181. URL: <https://doi.org/10.15407/pp2020.02-03.173>.
4. Бармак О. В., Манзюк Е. А., Крак Ю. В., Куляс А. І. Принципи та підходи до формування ансамблів класифікаторів на підставі агрегування їх результатів. *Штучний інтелект*. 2018. Вип. 81, №3. С. 62–69.
5. Manziuk E. Intelligent information technology for obtaining trust decisions based on the ontology of trust in a human-centered approach. *Computer Systems And Information Technologies*. 2022. №1. Pp. 83–88. URL: <https://doi.org/10.31891/CSIT-2022-1-11>.
6. Бармак О.В., Калита О.Д., Манзюк Е.А. Модель для розпізнавання мімічних проявів емоцій. *Вісник Хмельницького національного університету. Серія «Технічні науки»*. 2020. Вип. 281, №1. С. 25–29. URL: <https://doi.org/10.31891/2307-5732-2020-281-1-25-29>.
7. Гримак Р.Р., Пасічник О.А., Скрипник Т.К., Манзюк Е.А. Інформаційна технологія прийняття контрольованих критично-безпекових рішень перетворення параметрів моделей при трансфері між системами візуалізації. *Вісник Хмельницького національного університету. Серія «Технічні науки»*. 2021. Вип. 299, №4. С. 35–42. URL: <https://doi.org/10.31891/2307-5732-2021-297-3-35-42>.
8. Джурабаєв О. В., Бармак О. В., Манзюк Е. А., Скрипник Т. К. Інформаційна зосередженість змістовності в тексті. *Вісник Хмельницького національного університету. Серія «Технічні науки»*. 2019. Вип. 275, №4. С. 80–83. URL: <https://doi.org/10.31891/2307-5732-2019-275-4-80-83>.
9. Манзюк Е. Метод паралельного формування кортежів ознак даних сегменту складнокласифікованої інформації в умовах невизначеності прийняття рішень та аналізу впливу на розмежованість інформації. *Вісник Хмельницького національного університету. Серія «Технічні науки»*. 2021. Вип. 297, №3. С. 232–238. URL: <https://doi.org/10.31891/2307-5732-2021-297-3-232-238>.
10. Скрипник Т. К., Манзюк Е. А., Москалюк С. С. Ефективний інтерфейс користувача. *Вісник Хмельницького національного університету. Серія «Технічні науки»*. 2017. Вип. 245, №1. С. 132–135.
11. Скрипник Т. К., Манзюк Е. А., Свистун С. О. Архітектура каркасу model-view-controller при розробці веб-додатків. *Вісник Хмельницького національного університету. Серія «Технічні науки»*. 2017. Вип. 249, №3. С. 208–212.
12. Манзюк Е. Комплексний метод визначення відповідності онтології довіри до інтелектуальних інформаційних технологій та структурованого домену.

Вимірювальна та обчислювальна техніка в технологічних процесах. 2021. №2. С. 103–108. URL: <https://doi.org/10.31891/2219-9365-2021-68-2-13>.

Статті у періодичних виданнях, включених до категорії «А» Переліку наукових фахових видань України, або у закордонних виданнях, проіндексованих у базах даних *Web of Science Core Collection* та/або *Scopus*:

13. Barmak A. V., Krak Y. V., Manziuk E. A., Kasianiuk V. S. Information technology of separating hyperplanes synthesis for linear classifiers. *Journal of Automation and Information Sciences*, 2019. Vol. 51, No. 5. Pp. 54–64. URL: <https://doi.org/10.1615/JAutomatInfScien.v51.i5.50>. (індексована в наукометричній базі *Scopus*, кuartиль Q3 відповідно до класифікації *SCImago Journal and Country Rank*)

14. Krak I., Barmak O., Manziuk E. Using visual analytics to develop human and machine-centric models: A review of approaches and proposed information technology. *Computational Intelligence*, 2022. Vol. 38, No. 3. Pp. 921–946. URL: <https://doi.org/10.1111/coin.12289>. (індексована в наукометричній базі *Scopus* та *Web of Science Core Collection*, кuartиль Q2 відповідно до класифікації *SCImago Journal and Country Rank*)

15. Krak Iu. V., Kasianiuk V. S., Kudin H. I., Barmak O. V., Manziuk E. A. Multivariate scaling of the characteristic features based on pseudo-inverse operations for recognition problems solving. *Pattern Recognition and Image Analysis*, 2020. Vol. 30, No. 2. Pp. 184–191. URL: <https://doi.org/10.1134/S1054661820020078>. (індексована в наукометричній базі *Scopus* та *Web of Science Core Collection*, кuartиль Q3 відповідно до класифікації *SCImago Journal and Country Rank*)

16. Manziuk E. A., Barmak A. V., Krak Y. V., Kasianiuk V. S. Definition of information core for documents classification. *Journal of Automation and Information Sciences*, 2018. Vol. 50, No. 4. Pp. 25–34. URL: <https://doi.org/10.1615/JAutomatInfScien.v50.i4.30>. (індексована в наукометричній базі *Scopus*, кuartиль Q3 відповідно до класифікації *SCImago Journal and Country Rank*)

17. Manziuk E. A., Wójcik W., Barmak O. V., Krak I. V., Kulias A. I., Drabovska V. A., Puhach V. M., Sundetov S., Mussabekova A. Approach to creating an ensemble on a hierarchy of clusters using model decisions correlation. *Przegląd Elektrotechniczny*, 2020. Vol. 96, No. 9. Pp. 108–113. URL: <https://doi.org/10.15199/48.2020.09.23>. (індексована в наукометричній базі *Scopus* та *Web of Science Core Collection*, кuartиль Q3 відповідно до класифікації *SCImago Journal and Country Rank*)

Монографії (розділи у колективних монографіях):

18. Mazurets O., Barmak O., Krak I., Manziuk E., Bahrii R. Method for adaptive semantic testing of educational materials level of knowledge. *Lecture Notes on Data Engineering and Communications Technologies*. 2022. Vol. 77. Pp. 491–506. https://doi.org/10.1007/978-3-030-82014-5_33. (ISBN 978-3-030-82013-8, індексована в наукометричній базі *Scopus*)

19. Krak Iu., Kruchynin K., Barmak O., Manziuk E., Kruchinin S.P. Visual analytics in machine training systems for effective decision. *Advanced Nanomaterials for Detection of CBRN. NATO Science for Peace and Security Series A: Chemistry and Biology*. 2020. Pp. 327–338. URL: https://doi.org/10.1007/978-94-024-2030-2_25. (ISBN 978-94-024-2029-6, індексована в наукометричній базі *Scopus*)

Публікації, які засвідчують апробацію матеріалів дисертації:

20. Barmak O., Kalyta O., Krak I., Manziuk E., Kuznetsov V. Model of the facial emotions expressions based on grouping classes of feature vectors. *Lecture Notes in Computational Intelligence and Decision Making*. 2021. Vol. 1246. Pp. 65–76. URL: https://doi.org/10.1007/978-3-030-54215-3_5. (індексована в наукометричній базі Scopus)
21. Barmak O., Manziuk E., Krak I. Classification based hierarchical clustering prediction variability in the ensembles of models using a statistical Approach. *The 2020 IEEE 15th International Scientific and Technical Conference “Computer Science and Information Technologies”*: Proceedings (Lviv-Zbarazh (Ukraine), September 23-26, 2020). Vol. 1. Pp. 11–14. URL: <https://doi.org/10.1109/CSIT49958.2020.9322019>. (індексована в наукометричній базі Scopus)
22. Barmak O., Kalyta O., Manziuk E., Krak I. Simplified model for recognition facial emotions. *The 2019 IEEE International Scientific and Technical Conference “Advanced Trends in Information Theory”*: Proceedings (Kyiv (Ukraine), December 18-20, 2019). Pp.492–495. URL: <https://doi.org/10.1109/ATIT49449.2019.9030516>. (індексована в наукометричній базі Scopus)
23. Manziuk E., Barmak O., Krak I., Mazurets O., Pylypiak O. Method of features analysis on transition data. *The 2021 IEEE International Scientific and Technical Conference “Advanced Trends in Information Theory”*: Proceedings (Kyiv (Ukraine), December 15-17, 2021). Pp. 272–277. URL: <https://doi.org/10.1109/ATIT54053.2021.9678787>. (індексована в наукометричній базі Scopus)
24. Barmak O., Manziuk E., Krak I. Using piecewise hyper linear classification in multidimensional feature space for text content. *The 2019 IEEE 14th International Scientific and Technical Conference “Computer Sciences and Information Technologies”*: Proceedings (Lviv (Ukraine), September 17-20, 2019). Pp. 119–123. URL: <https://doi.org/10.1109/STC-CSIT.2019.8929798>. (індексована в наукометричній базі Scopus)
25. Krak I., Barmak O., Manziuk E., Kudin H. Approach to piecewise-linear classification in a multi-dimensional space of features based on plane visualization. *Lecture Notes in Computational Intelligence and Decision Making*. 2020. Vol. 1020. Pp. 35–47. URL: https://doi.org/10.1007/978-3-030-26474-1_3. (індексована в наукометричній базі Scopus та Web of Science Core Collection)
26. Krak I., Barmak O., Manziuk E., Kulias A. Data classification based on the features reduction and piecewise linear separation. *Intelligent Computing and Optimization*. 2020. Vol. 1072. Pp. 282–289. URL: https://doi.org/10.1007/978-3-030-33585-4_28. (індексована в наукометричній базі Scopus)
27. Barmak O., Krak I., Manziuk E., Lytvynenko V., Kalyta O. Classification technology based on hyperplanes for visual analytics with implementations for different subject areas. *CEUR-WS*. 2020. Vol. 2623. Pp. 96–106. (індексована в наукометричній базі Scopus та Web of Science Core Collection)
28. Barmak O., Krak I., Manziuk E. Diversity as the basis for effective clustering-based classification. *CEUR-WS*. 2020. Vol. 2711. Pp. 53–67. (індексована в наукометричній базі Scopus)
29. Krak I., Petrovych V., Kuznetsov V., Manziuk E., Barmak O., Kulias A. On classification hidden concepts language in specialized texts based on methods of the intellectual data processing. *CEUR-WS*. 2021. Vol. 2864. Pp. 110–120. (індексована в наукометричній базі Scopus)

30. Manziuk E., Barmak O., Krak I., Mazurets O., Skrypnyk T. Formal model of trustworthy artificial intelligence based on standardization. *CEUR-WS*. 2021. Vol. 2853. Pp. 190–197. (індексована в наукометричній базі *Scopus*)
31. Manziuk E., Krak I., Barmak O., Mazurets O., Kuznetsov V., Pylypiak O. Structural alignment method of conceptual categories of ontology and formalized domain. *CEUR-WS*. 2021. Vol. 3003. Pp. 11–22. (індексована в наукометричній базі *Scopus*)
32. Petrovych V., Kuznetsov V., Manziuk E., Krak I., Kasianiuk V., Barmak O., Kulias A. I. On development classification methods for hidden features separation in data. *CEUR-WS*. 2021. Vol. 3018. Pp. 25–31. (індексована в наукометричній базі *Scopus*)
33. Barmak O., Manziuk E., Kalyta O., Krak I., Kuznetsov V., Kulias A. Recognition of emotional expressions using the grouping crowdings of characteristic mimic states. *CEUR-WS*. 2020. Vol. 2866. Pp. 173–181. (індексована в наукометричній базі *Scopus*)
34. Barmak O., Krak Y., Manziuk E. Characteristics for choice of models in the ansables classification *CEUR-WS*. 2018. Vol. 2139. Pp. 171–179. (індексована в наукометричній базі *Scopus*).
35. Krak I. V., Kudin H. I., Barmak O. V., Manziuk E. O., Smolarz A., Mamyrbaev O. Method and algorithm of the piecewise-hyperplane clusterization using tools of pseudo-inverse matrices. *SPIE - The International Society for Optical Engineering “Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019”*: Proceedings (Wilga (Poland), November 06, 2019). Vol. 11176. Pp.1–8. URL: <https://doi.org/10.1117/12.2537417>. (індексована в наукометричній базі *Scopus* та *Web of Science Core Collection*)
36. Krak I. V., Barmak O. V., Manziuk E. Visual analytics to build a machine learning model. *Research Advancements in Smart Technology, Optimization, and Renewable Energy*. IGI Global, 2021. Pp. 313-329. <https://doi.org/10.4018/978-1-7998-3970-5.ch015>
37. Krak Iu., Amirgaliev E., Wojcik W., Barmak O., Manziuk E. Information technology for classifying based on the methods of features data reduction and piecewise linear separation. *Информатика и прикладная математика*, матеріали міжнар. наук.-прак. конф. (Алмати (Казахстан), 25-29 вересня 2019 р.). Алматы, 2019. С. 339–350.
38. Krak Iu., Barmak O., Manziuk E., Kasianiuk V. Ensembles for classification base on their aggregation results. *Problem of decision making under uncertainties*. тези доп. міжнар. наук.-прак. конф. (Хургада (Египет), 24 січня – 1 лютого 2019). Київ, 2019. С. 52–53.
39. Бармак А. В., Крак Ю. В., Манзюк Е. А., Мазурець О.В., Пилип'як О.В. Визначення нерелевантних ознак в малих даних на базі кластерного підходу. *Інформаційні управляючі системи і технології*: матеріали міжнар. наук.-практ. конф. (Одеса, 23–25 вересня 2021 р.). Одеса, 2021. С. 96–99.
40. Бармак А.В., Крак Ю. В., Манзюк Э. А. Разнообразие как основа эффективной классификации на базе кластеризации. *Інформаційні управляючі системи і технології*: матеріали міжнар. наук.-практ. конф. (Одеса, 24–26 вересня 2020 р.). Одеса, 2020. С.71–73.
41. Бармак О.В., Касьянюк В.С., Крак Ю.В., Манзюк Е.А. Класифікація інформації на основі методу візуалізації даних та машинного навчання. *Ідеї академіка В.М. Глушкова і сучасні проблеми штучного інтелекту: Глушковські читання*: матеріали всеукр. наук.-прак. конф. (Київ, 29 листопада 2019 р.). Київ, 2019. С. 15–16.

42. Бармак О.В., Манзюк Е.А., Крак Ю.В., Кудін Г.І., Куляс А.І. Аналіз та візуалізація даних в системах людино-машинного навчання. *Математичне та програмне забезпечення інтелектуальних систем*: тези доп. міжнар. наук.-прак. конф. (Дніпро, 18-20 листопада 2020 р.). Дніпро, 2020. С. 29–30.

43. Бородін М. Ю., Манзюк Е.А., Скрипник Т.К. Забезпечення захищеності програмних систем з використанням трансформаційних перетворювань. *Актуальні проблеми комп'ютерних наук*: збірник наукових праць за матеріалами всеукр. наук.-прак. конф. (Хмельницький, 14-15 листопада 2019 р.). Хмельницький, 2019. С. 35–38.

44. Тростинський Н.М., Пасічник О.А., Скрипник Т.К., Манзюк Е.А. Інформаційна технологія оптимізації контрольованих людиною критично-безпекових рішень при перегляді хмар точок “web point cloud viewer.” *Вісник Хмельницького національного університету. Серія «Технічні науки»*. 2021. Вип. 299, №4. С. 11–17. URL: <https://doi.org/10.31891/2307-5732-2021-297-3-11-17>.

45. Гордійчук Б. Г., Манзюк Е. А., Скрипник Т. К. Виявлення аномалій в даних. *Актуальні проблеми комп'ютерних наук*: збірник наукових праць за матеріалами всеукр. наук.-прак. конф. (Хмельницький, 9-10 листопада 2020 р.). Хмельницький, 2020. С. 72–74.

46. Джурабаєв О.В., Бармак О.В., Манзюк Е.А. Пошук змісту в текстовій інформації. *Інтелектуальний потенціал - 2018*: збірник наукових праць за матеріалами всеукр. наук.-прак. конф. (Хмельницький, 14-16 листопада 2018 р.). Хмельницький, 2018. С. 35–39.

47. Крак Ю.В., Касьянюк В.С., Бармак О.В., Манзюк Е.А. Побудова моделі класифікації даних для багатовимірного простору ознак. *Математичне та програмне забезпечення інтелектуальних систем*: тези доп. міжнар. наук.-прак. конф. (Дніпро, 20-22 листопада 2019 р.). Дніпро, 2019. С. 149–150.

48. Манзюк Е. А., Скрипник Т. К. Система цільової кластеризації на послідових даних. *Актуальні проблеми комп'ютерних наук*: збірник наукових праць за матеріалами всеукр. наук.-прак. конф. (Хмельницький, 15-16 жовтня 2021 р.). Хмельницький, 2021. С. 364–366.

49. Манзюк Е. А., Бармак О. В., Крак Ю. В. Про підходи до машинного навчання на основі довіри штучному інтелекту. *Історія, сучасний стан та тенденції цифрового розвитку суспільства: Глушковські читання*: матеріали всеукр. наук.-прак. конф. (Київ, 2 грудня 2021 р.). Київ, 2021. С. 13–15.

50. Манзюк Е.А., Бармак О.В., Крак Ю.В., Касьянюк В.С. Використання комп'ютерного навчання та візуалізації даних для ефективної класифікації інформації. *Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту*: збірник наукових праць за матеріалами міжнар. наук. конф. (Залізний Порт, 25-29 травня 2020 р.). Залізний Порт, 2020. С. 86.

51. Манзюк Е.А., Крак Ю.В., Бармак О.В. Методи машинного навчання на основі довіри штучному інтелекту. *Математичне та програмне забезпечення інтелектуальних систем*: тези доп. міжнар. наук.-прак. конф. (Дніпро, 17-19 листопада 2021 р.). Дніпро, 2021. С. 130–131.

52. Манзюк Е. А., Скрипник Т. К., Гірний М. Ю. Використання штучного інтелекту для розпізнавання складових елементів об'єктів на базі зображення. *Computer Systems And Information Technologies*. 2020. №1. С. 41–45. URL: <https://doi.org/10.31891/CSIT-2020-1-5>.

53. Манзюк Е.А., Крак Ю.В., Бармак О.В. Про розробку методів машинного навчання на основі довіри штучному інтелекту. *Актуальні проблеми теорії*

керуючих систем у комп'ютерних науках: праці наук.-практ. конф. (Слов'янськ, 21-24 грудня 2021 р.). Слов'янськ, 2021. С. 70–72.

54. Прокопов Р. І., Манзюк Е. А., Скрипник Т. К. Інформаційна система для визначення подібності документів. *Актуальні проблеми комп'ютерних наук: збірник наукових праць за матеріалами всеукр. наук.-практ. конф. (Хмельницький, 9-10 листопада 2020 р.). Хмельницький, 2020. С. 232–236.*

55. Скрипник Т., Манзюк Е. Аналітична система визначення якості перекладу текстової інформації методами машинного навчання. *Українсько-Польські наукові діалоги: тези доп. міжнар. конф. (Хмельницький – Кам'янець-Подільський, 20-23 жовтня 2021). Хмельницький – Кам'янець-Подільський, 2021. С. 151–153.*

56. Побережний П. В., Манзюк Е. А., Скрипник Т. К. Інформаційна система класифікації текстової інформації. *Сучасні технології в механіці: збірник наукових праць міжнар. конф. (Хмельницький, 19-21 квітня 2018 р.). Хмельницький, 2018. С. 169-173.*

57. Кузьмінський М. С., Манзюк Е. А. Система прогнозування продажів сервісних послуг в системах обслуговування. *Актуальні проблеми комп'ютерних наук: збірник наукових праць за матеріалами всеукр. наук.-практ. конф. (Хмельницький, 9-10 листопада 2020 р.). Хмельницький, 2020. С. 157–158.*

58. Бармак О.В., Крак Ю.В., Мазурець О.В., Манзюк Е.А. Ментально-формальні рішення машинного навчання для інформаційної технології автоматизованого створення тестів у сфері безпеки та медицини. *Інформаційні управляючі системи і технології: матеріали міжнар. наук.-практ. конф. (Одеса, 23–25 вересня 2021 р.). Одеса, 2021. С. 105-107.*

59. Скрипник Т. К., Манзюк Е. А. Метод машинного навчання для визначення якості перекладу текстової інформації. *Актуальні проблеми комп'ютерних наук: збірник наукових праць за матеріалами всеукр. наук.-практ. конф. (Хмельницький, 15-16 листопада 2021 р.). Хмельницький, 2021. С. 404–405.*

Публікації, які додатково відображають наукові результати дисертації:

60. Літературний письмовий твір наукового характеру «Інформаційна технологія раннього діагностування захворювань легень за ментально-формальною нейромережевою моделлю»/ П.М. Радюк, О.В. Бармак, Е.А. Манзюк, О.В. Мазурець, Р.О. Багрій, М.О. Молчанова, О.В. Собко. – Свідоцтво про реєстрацію авторського права на твір № 109538 від 17.11.21 р., Державний департамент інтелектуальної власності.

61. Літературний письмовий твір наукового характеру «Інформаційна система візуального подання та пояснення результатів ідентифікації захворювання легень за медичними зображеннями грудної клітини»/ П.М. Радюк, О.В. Бармак Е.А. Манзюк, О.В. Мазурець, Р.О. Багрій, М.О. Молчанова, О.В. Собко. – Свідоцтво про реєстрацію авторського права на твір № 109539 від 17.11.21 р., Державний департамент інтелектуальної власності.

62. Твір наукового характеру «Інформаційна система альтернативної комунікації для безконтактного введення текстових повідомлень з використанням обмеженої кількості простих рухів»/ Р.О. Багрій, О.В. Бармак, О.В. Мазурець, Е.А. Манзюк, П.М. Радюк, М.О. Молчанова, О.В. Собко. – Свідоцтво про реєстрацію авторського права на твір № 109262 від 09.11.21 р., Державний департамент інтелектуальної власності.