

Краснюк С. О., ст. викл.

Київський національний університет технологій та дизайну

Краснюк М. Т., к. е. н., доцент,

Київський національний економічний університет імені Вадима Гетьмана

**АКТУАЛЬНІСТЬ НАВЧАЛЬНОГО КУРСУ
„DATA MINING in BIG DATA” З УРАХУВАННЯМ
ОСТАННІХ ГЛОБАЛЬНИХ ЕКОНОМІЧНИХ ТА ОСВІТНІХ
ТЕНДЕНЦІЙ В ЄС ТА США**

Всі підприємства (торгівельні, виробничі, наукові, сервісні тощо) реєструють, записують та зберігають величезні обсяги різномірної інформації (кількісної, якісної, текстової, мультимедійної тощо) щодо всіх аспектів своєї діяльності. Причому швидкість та деталізація таких потоків зростають. Зокрема, опитування Європейських компаній показало, що майже половина з них очікують щорічне зростання потоків своїх даних на 25% .

Аналіз сучасної міжнародної практики корпоративного менеджменту показує наступні глобальні фактори, які мають значний і постійний вплив на інформаційний менеджмент:

- надвеликий простір рішень та їх мультидисциплінарність (при прийнятті управлінських рішень наростає кількість різноформатних вхідних змінних, їх комбінацій, та відповідних варіантів складених та невірноважених сценаріїв);

- поява та вдосконалення інноваційних ІТ-технологій (що, зокрема, передбачають тотальні можливості для реєстрації, передавання, зберігання, обробки та автоматизованого аналізу всіх даних, подій та станів);

- всеосяжна глобалізація, on-line конкуренція та ускладнення структури ринку міжнародних мультимодальних послуг/товарів призводять до появи нових та зростання впливу відомих чинників невизначеності та неповноти даних, збільшення «інформаційного шуму»;

- зростання динаміки флуктуацій суб'єктивних та об'єктивних факторів, збільшення частки аномалій в даних, зміна закономірностей їх

взаємодії, зміна значущості факторів у вже побудованих моделях (при чому постає нагальна потреба не тільки автоматично реагувати на ці динаміки в режимі реального часу, але і автоматизувати верифікацію та перенавчання існуючих моделей);

- радикальне зростання потоків неструктурованих даних всіх типів від автономних пристроїв, сенсорів та датчиків.

Не зважаючи, на викладені вище перепони, які мають значний і постійний вплив на управлінську діяльність в сфері міжнародного корпоративного менеджменту, проведений аналіз і діагностика вітчизняної практики менеджменту, показали, що системною проблемою для вітчизняного корпоративного менеджменту все ще є відсутність цілісної формалізованої моделі управління бізнес-процесами і, як наслідок, суб'єктивізм і інтуїтивне (ручне) управління. З огляду на вищевказане, можна деталізувати додаткові актуальні проблеми антикризового управління українських корпорацій:

- відсутність формалізованих описів бізнес-процесів та відповідних корпоративних стандартів (або відсутність автоматизації контролю їх виконання);

- «ручне» та епізодичне управління якістю послуг/товарів та лояльністю клієнтів;

- значний вплив на показники корпорацій некомпетентності/суб'єктивності/зловживань з боку операційного менеджменту;

- відсутність автоматизації та об'єктивності при прогнозуванні попиту, а, отже, і падіння ефективності операційної діяльності;

- епізодичне та часткове використання всіх накопичених та/або доступних внутрішніх та зовнішніх даних (особливо потокових даних в режимі «реального часу»), слабкий або відсутній вплив результатів аналізу вищезгаданих даних не тільки на операційному рівні, але і на рівні тактичного і стратегічного менеджменту.

Отже, крім вищезазначеного впливу Big Data, ще одним фактором, що ускладнює аналітику сучасних корпоративних даних, є те, що більшість даних, збережених в DBMS, а тим більше потокових - мають характер просторово-часових рядів, отже актуальна потреба не тільки в класичному статистичному аналізі багатомірних часових рядів, але і потреба в пошуку прихованих невідомих закономірностей та подальшій побудові прогнозних просторових моделей.

Тобто, з урахуванням дослідженої вище специфіки сучасного менеджменту, можна стверджувати, що класичного функціоналу регресійного, факторного, дисперсійного аналізу в сучасних кризових умовах для вітчизняних компаній – не достатньо .

Адже для того, щоб сформулювати нові передумови для ухвалення ефективних управлінських (в т.ч. антикризових) рішень, потрібні нові, об'єктивні знання про приховані суті, зв'язки і закономірності досліджуваної предметної області. Відшукати та перевірити (в т.ч. в ході класичного статистичного аналізу) актуальні факти та гіпотези в базі чи сховищі даних не так вже складно, проте в сучасній конкурентній інформаційній економіці потрібні не просто факти, а нові, об'єктивні закономірності, верифіковані, інтерпретовані та формалізовані у відповідних моделях подання знань .

В якості основного джерела таких нових об'єктивних управлінських шаблонів для компанії, яка діє на ризикованому, корупційному, конкурентному та кризовому вітчизняному ринку - може бути лише технологія інтелектуального аналізу даних (Data Mining).

Data Mining - це виявлення в первинних, накопичуваних в результаті оброблення бізнесових транзакцій структурованих та неструктурованих даних, наперед невідомих чи прихованих закономірностей і шаблонів (знань) з метою формування відповідної бази знань та прийняття обґрунтованих і оптимальних бізнесових рішень.

Можна стверджувати, що в сучасних умовах розвитку глобальної економіки, і в зв'язку з появою нових галузей економічної діяльності, гібридне застосування технологій Data Mining та Big Data зумовлює отримання додаткових конкурентних переваг, а отже і підвищення інвестиційної привабливості і капіталізації.

Зважаючи на більш ніж двадцятирічний період науково-практичного розвитку Data Mining, і, враховуючи те, що в провідних університетах США та ЄС, при підготовці MBA та MBI широко розповсюджене не тільки вивчення відповідних навчальних дисциплін, пов'язаних з концепцією Data Mining, але і запроваджено окремі магістерські програми щодо Data Science — для сучасних конкурентних кваліфікованих молодих українських випускників ЗВО спеціальностей менеджмент, фінанси - беззаперечно актуальність мають відповідні компетенції щодо всього спектру алгоритмів Data Mining та технології Big Data. *На доказ актуальності вищенаведеного свідчать такі дані, станом на 28 квітня 2019 року:*

- на провідному вітчизняному кадровому ресурсі Work.ua кількість вакансій data analyst по Україні становить **68** осіб, а вакансій data scientist становить **563** вакансії;

- на провідному професійному ресурсі в галузі Data Mining - <https://www.kdnuggets.com/education/usa-canada.html> пропонується **120** діючих в USA/Canada бакалаврських та магістерських спеціалізованих Degree Programs, присвячених Education in Analytics, Data Science, Machine Learning;

- на провідному професійному ресурсі в галузі Data Mining - <https://www.kdnuggets.com/education/europe.html> пропонується **67** діючих в ЄС бакалаврських та магістерських спеціалізованих Degree Programs, присвячених Education in Analytics, Data Science, Machine Learning;

- на провідному ресурсі on-line освіти <https://ru.coursera.org/courses?query=data%20science&> у відповідному розділі наразі пропонується 905 активних on-line пропозицій про навчання та отримання дипломів в сфері Data Science.

Зважаючи на вищенаведене, та на вимоги до інноваційної складової освітньої діяльності в рамках Болонського процесу, авторами обґрунтовано рекомендується запровадити вивчення навчальної дисципліни „DATA MINING in BIG DATA” в робочі навчальні плани на бакалаврському рівні вищої освіти по спеціальностям «Економіка», «Менеджмент», «Фінанси» для іноземних та вітчизняних студентів.

Мета вищезазначеної дисципліни – сформувати поглиблені теоретичні знання та ґрунтовні практичні навички щодо ефективного застосування технології, алгоритмів та засобів інтелектуального аналізу даних та знань.

Предмет дисципліни: ефективне застосування алгоритмів та технологій інтелектуального аналізу даних та знань, проектування, створення та налаштування модулів аналізу даних та знань в економічних інформаційних системах.

Data Mining - мультидисциплінарна наука, яка передбачає поєднання широкого математичного інструментарію і останніх досягнень в сфері інформаційних технологій. У технології Data Mining гармонійно об'єдналися формалізовані методи і методи неформалізованого аналізу, методи врахування невизначеності в даних та знаннях.

Завдання вищезазначеної дисципліни:

- вивчення методологічно-організаційних особливостей прийняття управлінських рішень на основі інтелектуального аналізу даних та знань;
- засвоєння специфіки алгоритмів та підходів до інтелектуального аналізу даних та знань в інформаційних системах;
- опрацювання передового досвіду використання алгоритмів та систем інтелектуального аналізу даних та знань в галузях корпоративного планування, менеджменту, маркетингу, фінансово-кредитної системи, прогнозування тощо;
- отримання практичних навичок ефективного використання і адаптації всього спектру систем інтелектуального аналізу даних та знань;
- набуття уміння проектувати, розробляти, впроваджувати та ефективно супроводжувати системи інтелектуального аналізу даних та знань різного

призначення і різної проблемної орієнтації на всіх стадіях життєвого циклу корпоративних інформаційних систем.