

від початкових даних. Випадковий ліс стійкий до шуму, але ресурсозатратний у генерації і роботі. Метод опорних векторів дає максимально точний результат, але порівняно повільний у роботі. Вибір алгоритму повинен відповідати потребам – обсягу вхідних даних та вимогам користувача.

### Список використаних джерел

1. Machine learning explained. 2024. URL: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained> (дата звернення: 16.05.2024).
2. Optimization vs standartization. 2024. URL: <https://www.geeksforgeeks.org/normalization-vs-standardization/> (дата звернення: 16.05.2024).
3. Harmon M., Klabjan D. Activation Ensembles for Deep Neural Networks. arXiv, 2017. DOI: 10.48550/arXiv.1702.07790 (дата звернення: 16.05.2024).
4. Luo H., Haffner P., Paiement J.-F. Accelerated Parallel Optimization Methods for Large Scale Machine Learning. arXiv, 2014. DOI: 10.48550/arXiv.1411.6725 (дата звернення: 16.05.2024).
5. What is logistic regression. 2024. URL: <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-logistic-regression/> (дата звернення: 16.05.2024).

## УДК 519.2

*Мороз Д. В., здобувачка 1 курсу спеціальності 111 Математика, Луценко А. В., д-р філос. з математики, старший викладач кафедри прикладної математики та кібербезпеки*

### ЗАСТОСУВАННЯ ЛІНІЙНОЇ АЛГЕБРИ В ІНТЕЛЕКТУАЛЬНОМУ АНАЛІЗІ ДАНИХ

*Донецький національний університет імені Василя Стуса, м. Вінниця*

**Актуальність.** У світі, що переповнений даними, важливість їх аналізу та інтерпретації зростає з кожним днем. Інтелектуальний аналіз даних стає ключовим інструментом для виявлення закономірностей, отримання цінних інсайтів та прийняття стратегічних рішень у різних сферах, від бізнесу до медицини. Одним із фундаментальних математичних інструментів, який забезпечує основу для інтелектуального аналізу даних, є лінійна алгебра. Розуміння та ефективне використання принципів лінійної алгебри в цьому контексті стає важливим завданням для дослідників, аналітиків та фахівців з обробки даних.

**Метою роботи** є розгляд застосування лінійної алгебри в інтелектуальному аналізі даних. У роботі будуть розглянуті основні поняття лінійної алгебри, як-от вектори, матриці, лінійні простори та операції над ними, а також конкретні сфери застосування цих принципів у сучасному інтелектуальному аналізі даних, як-от обробка та аналіз даних, методи машинного навчання, обробка зображень та сигналів, а також рекомендаційні системи. Результатом роботи буде розуміння важливості лінійної алгебри для ефективного використання інтелектуальних методів аналізу даних та виявлення перспектив подальшого використання цих знань у практиці.

Основними поняттями лінійної алгебри є вектори, матриці, лінійні простори, операції додавання та множення [1].

**Застосування лінійної алгебри в інтелектуальному аналізі даних.** Представлення даних у вигляді матриць: дані можуть бути представлені у вигляді матриць, де кожен рядок відповідає окремому спостереженню, а кожен стовпець – означенню або функції. Цей підхід дає змогу ефективно зберігати та обробляти великі обсяги даних, а також виконувати різноманітні операції за допомогою матричних операцій [1].

*Зведення до матричних операцій:* багато методів обробки даних, як-от факторний аналіз, аналіз головних компонент та кластерний аналіз, можуть бути сформульовані у вигляді матричних операцій. Наприклад, у методі головних компонент використовуються операції знаходження власних значень та власних векторів матриці коваріації даних.

*Розв'язання систем лінійних рівнянь:* у деяких випадках обробка даних включає розв'язання систем лінійних рівнянь, що можуть виникнути під час застосування певних методів, наприклад, у задачах оптимізації або моделюванні. Лінійна алгебра надає ефективні методи для розв'язання таких систем, що дає змогу отримати числові рішення задач.

*Декомпозиція та розкладання матриць:* у деяких методах обробки даних використовуються операції декомпозиції та розкладання матриць, як-от SVD (Singular Value Decomposition) або QR-розклад. Ці операції дають змогу розкрити основні структурні властивості даних та зменшити їх розмірність, що може полегшити подальший аналіз.

*Визначення лінійних залежностей та кореляцій:* лінійна алгебра дає змогу визначити лінійні залежності та кореляції між різними змінними в наборі даних. Це допомагає виявляти структуру та взаємозв'язки між змінними, що може бути важливим під час розуміння даних та прийняття рішень на їх основі.

**Методи машинного навчання.** *Лінійна регресія* – це метод, який використовує лінійну функцію для прогнозування значення залежної змінної на основі набору незалежних змінних. У лінійній регресії кожен прихований шар моделі може бути представлений у вигляді вектора параметрів, а сама модель – у вигляді матриці ваг. Застосування лінійної алгебри дає змогу ефективно оцінювати параметри моделі та робити прогнози на основі нових даних [2].

*Метод головних компонент* – це метод зменшення розмірності даних, який знаходить лінійну комбінацію змінних, що максимізує дисперсію даних. У PCA використовуються матричні операції для знаходження власних значень та власних векторів коваріаційної матриці даних.

За допомогою лінійної алгебри виконуються операції зменшення розмірності та відновлення даних після зменшення розмірності.

*Метод опорних векторів* – це метод класифікації, який шукає оптимальну гіперплощину, що розділяє класи даних у просторі вищих вимірів. У SVM лінійна алгебра використовується для знаходження оптимальних ваг та зсуву, які визначають гіперплощину. Операції лінійної алгебри дають змогу ефективно знаходити оптимальні параметри моделі та робити прогнози для нових даних.

*Логістична регресія* – це метод класифікації, який використовує лінійну комбінацію змінних та логістичну функцію для передбачення ймовірності належності до певного класу. У логістичній регресії параметри моделі можуть бути оцінені за допомогою методів лінійної алгебри, наприклад, методом градієнтного спуску [3].

У багатьох архітектурах нейронних мереж, як-от багатошарові перцептрони, лінійна алгебра використовується для обчислення ваг та активаційних функцій. Операції лінійної алгебри виконуються для передачі сигналів між шарами та оновлення параметрів моделі під час навчання.

**Обробка зображень та сигналів з використанням лінійної алгебри.** Зображення та сигнали можуть бути представлені у вигляді матриць чисел, де кожен елемент відповідає інтенсивності пікселя або значенню сигналу. Це представлення дає змогу виконувати різноманітні операції обробки та аналізу за допомогою матричних операцій [4].

Лінійна алгебра використовується для застосування різних фільтрів до зображень, як-от розмивання, розкриття різкості, видалення шуму тощо. Фільтри часто представлені у вигляді матриць, які використовуються для зміни значень пікселів зображення. Лінійна алгебра використовується для зменшення обсягу зображень та сигналів шляхом видалення зайвої інформації або скорочення кількості бітів, необхідних для представлення даних. Методи компресії, як-от Singular Value Decomposition (SVD) або Wavelet Transform, використовують матричні операції для зменшення розміру даних. Після обробки або компресії зображень та сигналів лінійна алгебра може бути використана для відновлення початкових даних [5]. Матричні операції використовуються для виявлення взаємозв'язків та структурних особливостей у зображеннях та сигналах.

**Висновки.** Лінійна алгебра відіграє ключову роль в інтелектуальному аналізі даних, що є важливою галуззю у сучасному світі. Її застосування можливе від обробки та аналізу даних до розробки та вдосконалення методів машинного навчання. Завдяки лінійній алгебрі можливо ефективно представляти, обробляти та аналізувати складні набори даних, здійснювати класифікацію, прогнозування та рекомендації. Її розуміння є необхідним для фахівців у сфері аналізу даних та машинного навчання.

Використання лінійної алгебри в інтелектуальному аналізі даних є дуже перспективним.

#### Список використаних джерел

1. Strang G. Introduction to Linear Algebra, 2023.
2. Lang S. Linear Algebra. Springer, 2002.
3. Lay D., Lay S., McDonald J. Linear Algebra and Its Applications. Pearson, 2018.
4. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009.
5. Golub H., van Loan C. Matrix Computations. Johns Hopkins University Press, 2012.