

РОЗВІДУВАЛЬНИЙ АНАЛІЗ ДАНИХ ДЛЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ СТРЕСУ У СТУДЕНТІВ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ МЕТОДАМИ МАШИННОГО НАВЧАННЯ

Вінницький національний технічний університет

Анотація

Робота присвячена підготовці та розвідувальному аналізу даних для подальшого використання для інформаційної технології аналізу та прогнозування стресу у студентів закладів вищої освіти методами машинного навчання. Було проведено аналіз датасету та його ознак.

Ключові слова: розвідувальний аналіз даних, машинне навчання, інформаційна технологія, стрес, студенти, прогнозування, заклад вищої освіти.

Abstract

The work is devoted to the preparation and exploratory analysis of data for further use in the information technology of analysing and predicting stress in students of higher education institutions using machine learning methods. The dataset and its features were analysed.

Keywords: exploratory data analysis, machine learning, information technology, stress, students, prediction, higher education institution.

Вступ

Проблема стресу серед студентів закладів вищої освіти набуває особливої актуальності в умовах трансформації сучасної освітньої системи, що супроводжується інтенсивним навчальним навантаженням, соціально-економічною нестабільністю, інформаційним перевантаженням та активним впровадженням цифрових технологій. Регулярне перебування у стресових станах негативно позначається на психоемоційному стані студентської молоді, знижує рівень академічної мотивації, концентрацію уваги та загальну ефективність навчання [1,2].

Серед основних чинників, що провокують стрес у студентів, виділяють інтенсивне академічне навантаження, часті іспити, фінансову нестабільність, невизначеність щодо майбутнього, занижену самооцінку та соціальну ізоляцію. У деяких випадках стрес може виконувати роль стимулу до дії, однак здебільшого він призводить до емоційного виснаження, зниження працездатності та порушення психічної рівноваги.

У процесі навчання студенти часто втрачають мотивацію, що проявляється у зниженні зацікавленості до освітнього процесу, відмові від відвідування занять, пасивності у громадському житті закладу освіти, а також у зниженні академічних результатів. Стрес розглядається як негативна емоційна реакція на зовнішні чинники, хоча в помірних дозах він може бути важливим компонентом особистісного розвитку [3,4].

Вплив стресу проявляється у широкому спектрі психологічних і фізіологічних симптомів: порушення когнітивних функцій (концентрації, пам'яті), емоційна нестабільність, а також соматичні реакції організму. Усвідомлення складної природи стресу, а також впровадження сучасних підходів машинного навчання відкривають нові можливості для ідентифікації груп ризику та формування персоналізованих стратегій втручання. Такі стратегії спрямовані на підвищення рівня психологічної стійкості студентів і полегшення процесу їхньої адаптації до академічного середовища [5].

Актуальність цієї теми зумовлена необхідністю своєчасного виявлення та попередження негативних психоемоційних наслідків, які можуть мати довгостроковий вплив на якість освіти та добробут студентів. За умов воєнного стану, нестабільного інформаційного середовища та змін у форматах навчання (змішане, дистанційне), проблема набуває ще більшої гостроти.

У зв'язку з цим на перший план виходить потреба у впровадженні ефективних цифрових рішень для моніторингу, аналізу та прогнозування рівня стресу серед студентів. Інформаційні технології, зокрема алгоритми машинного навчання, дозволяють виявляти приховані взаємозв'язки між різними факторами стресу та формувати інтелектуальні системи підтримки прийняття рішень у сфері студентського добробуту.

Результати дослідження

Для проведення аналізу було обрано набір даних, що має назву «Student Stress Factors: A Comprehensive Analysis» та опублікований користувачем «Chhabii» та має відкритий доступ для загального використання на платформі Kaggle [6]. Даний датасет містить дані про фактори стресу 1100 студентів, а саме: психологічні фактори (рівень тривожності, самооцінка, історія психічного здоров'я, депресія), фізіологічні фактори (головний біль, кров'яний тиск, якість сну, проблеми з диханням), соціальні фактори (соціальна підтримка, тиск однолітків, позакласна діяльність, булінг), академічні фактори (академічна успішність, навчальне навантаження, стосунки викладача і студента, проблеми майбутньої кар'єри), фактори навколишнього середовища (рівень шуму, умови проживання, безпека, базові потреби). На рисунку 1 наведено фрагмент датасету «Student Stress Factors: A Comprehensive Analysis», що включає 21 ознаку.

	anxiety_level	self_esteem	mental_health_history	depression	headache	blood_pressure	sleep_quality	breathing_problem
0	14	20	0	11	2	1	2	4
1	15	8	1	15	5	3	1	4
2	12	18	1	14	2	1	2	2
3	16	12	1	15	4	3	1	3
4	16	28	0	7	2	3	5	1

Рис. 1. Приклад ознак, що містить набір даних

Даний датасет включає у себе такі змінні:

- Anxiety_level: рівень тривоги;
- Self_esteem: самооцінка;
- Mental_health_history: історія психічного здоров'я;
- Depression: депресія;
- Headache: головний біль;
- Blood_pressure: кров'яний тиск;
- Sleep_quality: якість сну;
- Breathing_problem: проблеми з диханням;
- Noise_level: рівень шуму;
- Living_conditions: умови проживання;
- Safety: безпека;
- Basic_needs: базові потреби;
- Academic_performance: успішність;
- Study_load: навчальне навантаження;
- Teacher_student_relationship: стосунки між викладачем та студентом;
- Future_career_concerns: майбутня кар'єра;
- Social_support: соціальна підтримка;
- Peer_pressure: тиск з боку однолітків;
- Extracurricular_activities: позакласна активність;
- Bullying: булінг;
- Stress_level: рівень стресу (цільова ознака).

Проведено попереднє очищення даних. Підготувавши дані створимо для них матрицю кореляції, що дасть змогу виявити явну залежність та зв'язки між різними параметрами та ознаками (рис.2) [7,8].

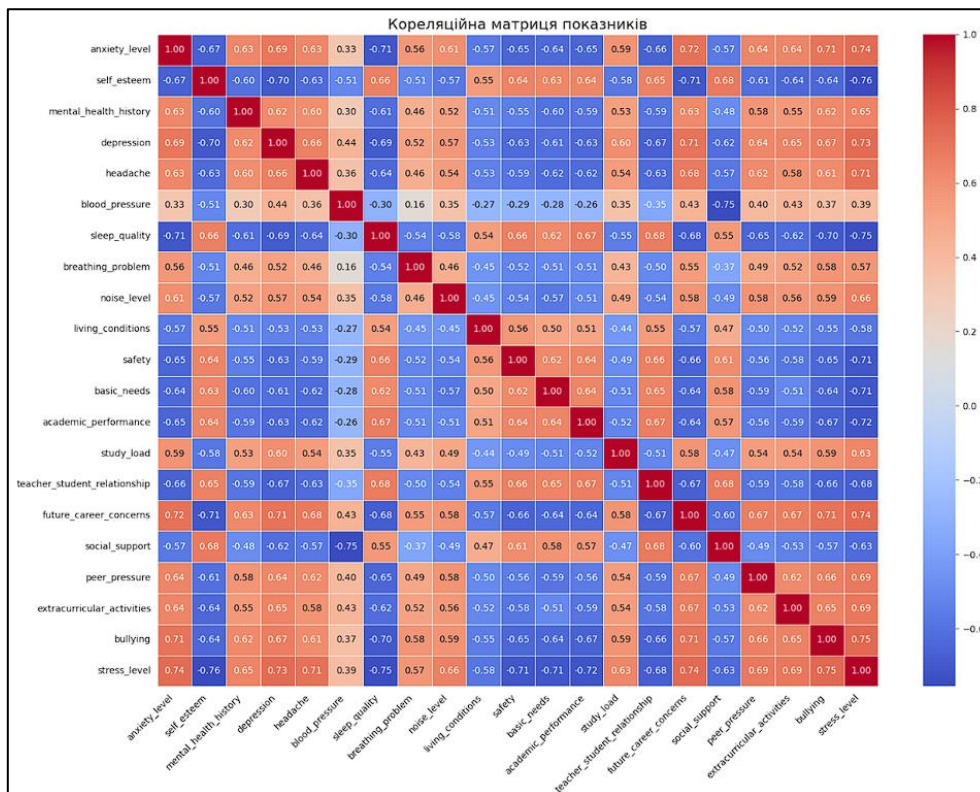


Рис. 2. Кореляційна матриця

Кореляційна матриця, представлена на рисунку 2, ілюструє взаємозв'язки між різними факторами життя студентів та рівнем стресу. Отже, найбільш впливовими факторами, що збільшують ймовірність виникнення стресу, є: рівень тривоги (Anxiety_level), депресія (Depression), булінг (Bullying), головний біль (Headache), історія психічного здоров'я (Mental_health_history), занепокоєння щодо майбутньої кар'єри (Future_career_concerns). У той час, такі аспекти, як висока самооцінка (Self_esteem), хороша академічна успішність (Academic_performance) та позитивні взаємини з викладачами (Teacher_student_relationship), понижують рівень стресу. Виявлені кореляції будуть використані для створення моделі машинного навчання, що допоможе точніше прогнозувати рівень стресу студентів.

Далі візуально відобразимо кількість наявних даних, відповідно до окремих параметрів (рис. 3) [9,10].

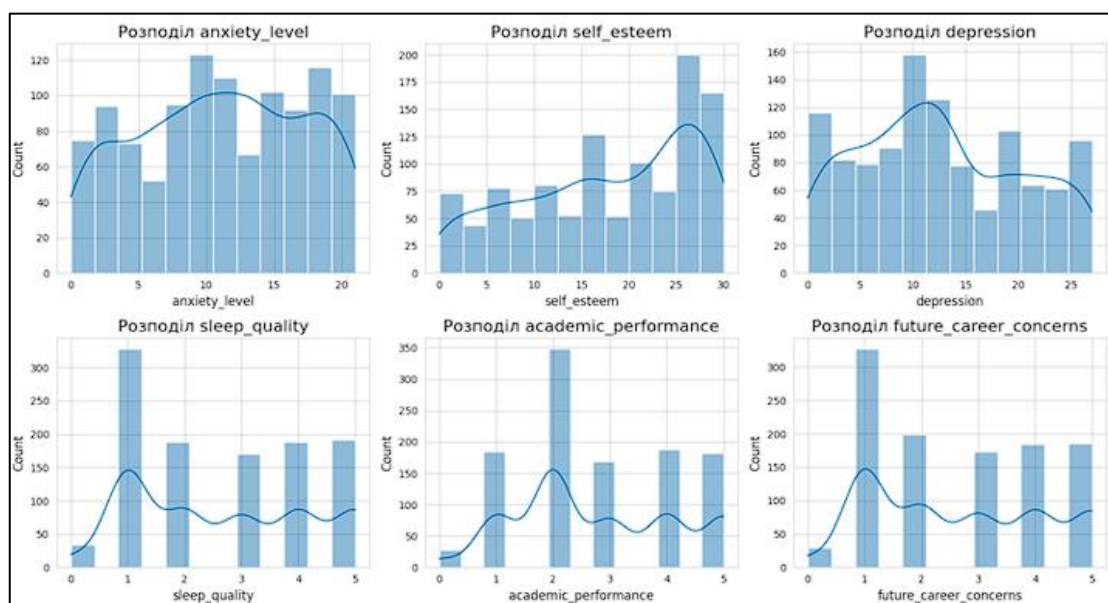


Рис. 3. Візуалізація кількості даних у вибірці по ознаках

Аналізуючи наведені графіки, ми можемо побачити, що значна частина студентів демонструє середній рівень тривожності. Однак, варто звернути увагу на те, що існує також суттєва кількість студентів, які мають високий показник тривоги. Це підкреслює важливість та актуальність проблем психічного здоров'я студентів. Також значна частина студентів має відносно високий рівень самооцінки. Але, наявність окремих груп з нижчою самооцінкою вказує на потребу в психологічній підтримці студентів. Аналізуючи результати візуалізації рівня депресії можемо бачити, що значна частина студентів, страждає від помірного рівня депресії, накладена крива, вказує на три основних максимумів, що говорить про різний рівень депресії студентів. На графіку також зображено дані рівня якості сну студентів, понад 300 студентів оцінюють свою якість сну як «1». Це вказує на те, що велика кількість студентів має проблеми зі сном. Згідно даними графіку, близько 350 студентів мають оцінку успішності «2», що свідчить про середній рівень академічної успішності. Цей показник є важливим, він демонструє, що студенти потребують додаткової підтримки та уваги з боку навчального закладу. У представленому графіку, близько 320 студентів оцінюють свій рівень занепокоєння майбутньою кар'єрою на рівні «2». Це свідчить про те, що певна кількість студентів демонструє мінімальний, але помітний ступінь тривоги щодо цього питання.

Далі проаналізуємо вплив окремих факторів на рівень стресу студентів. Ці візуалізації дають змогу оцінити, які саме змінні мають найбільшу вагу при класифікації рівня стресу та виявити закономірності їх взаємозв'язку.

На рисунку 4 зображено важливість ознак академічного чинника.

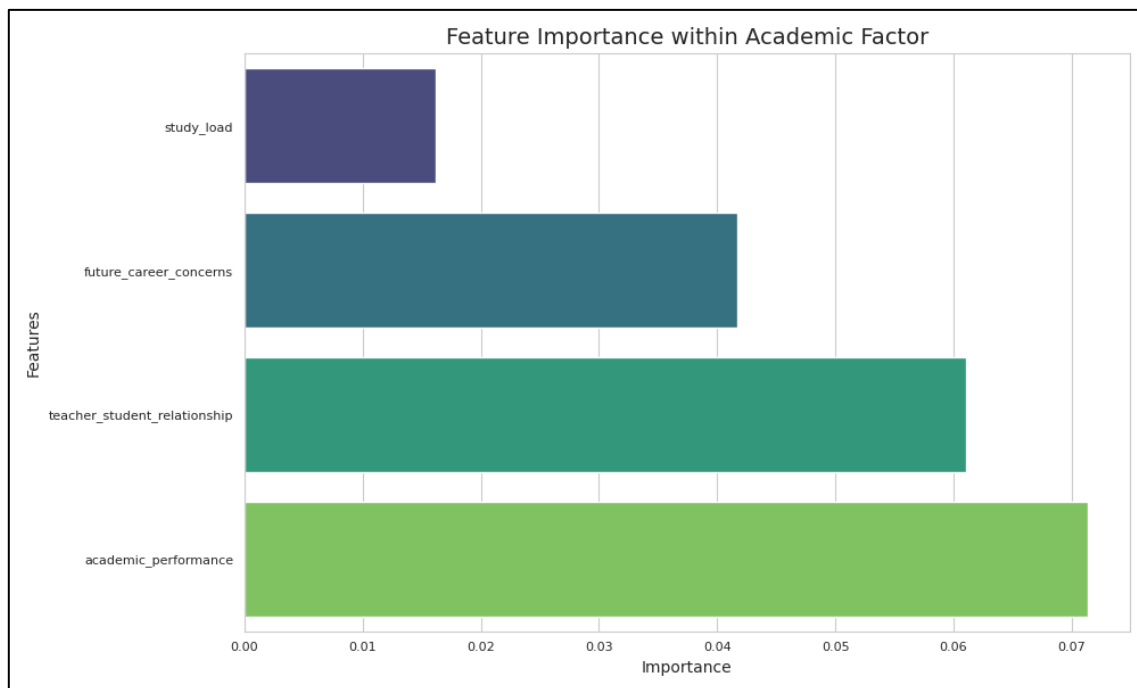


Рис. 4. Важливість ознак академічного чинника

Аналіз діаграми показує, що найбільшу вагу у структурі академічного чинника на рівень стресу має успішність (academic performance), це свідчить про те, що саме рівень навчальних досягнень студентів найбільше впливає на студентський стрес. На другому місці за значущістю перебувають стосунки між викладачем та студентом (teacher students relationship), що вказує на суттєвий вплив якості взаємин між викладачем і студентом на ймовірність виникнення стресу. Ознака майбутня кар'єра (future career concerns) має дещо нижчий рівень важливості, проте залишається значущим фактором. Найменш впливовою ознакою виявилось навчальне навантаження (study load), що може бути наслідком адаптації студентів до освітнього навантаження.

Проаналізуємо важливість ознак фізіологічних чинників, представлених на рисунку 5.

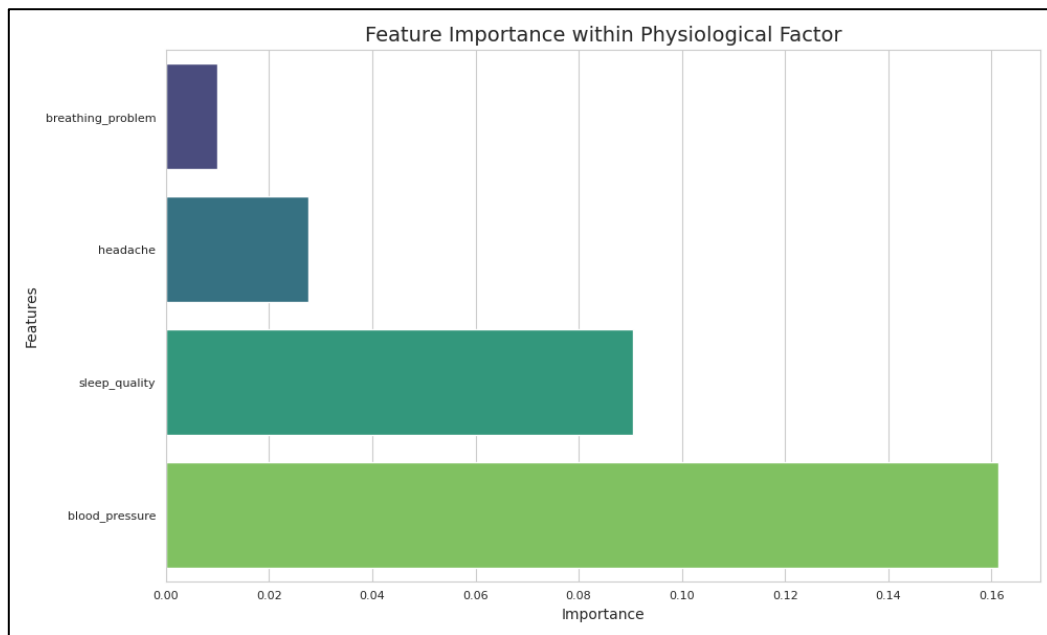


Рис.5. Важливість ознак фізіологічних чинників

Як видно з рисунка 5, кров'яний тиск (blood pressure) має найбільше значення важливості, близько 0,16. Це свідчить про те, що кров'яний тиск є найбільш значущою ознакою серед представлених на рівень стресу у студентів. Якість сну (sleep quality) посідає друге місце за важливістю, це вказує на значний внесок якості сну у фізіологічний стан студента. Головний біль (headache) має менше значення важливості, хоча головний біль і є важливим, його вплив значно менший порівняно з кров'яним тиском та якістю сну. Проблеми з диханням (breathing problem) демонструє найменше значення важливості серед усіх представлених.

Далі візуалізуємо важливість ознак чинника навколишнього середовища (рис.6).

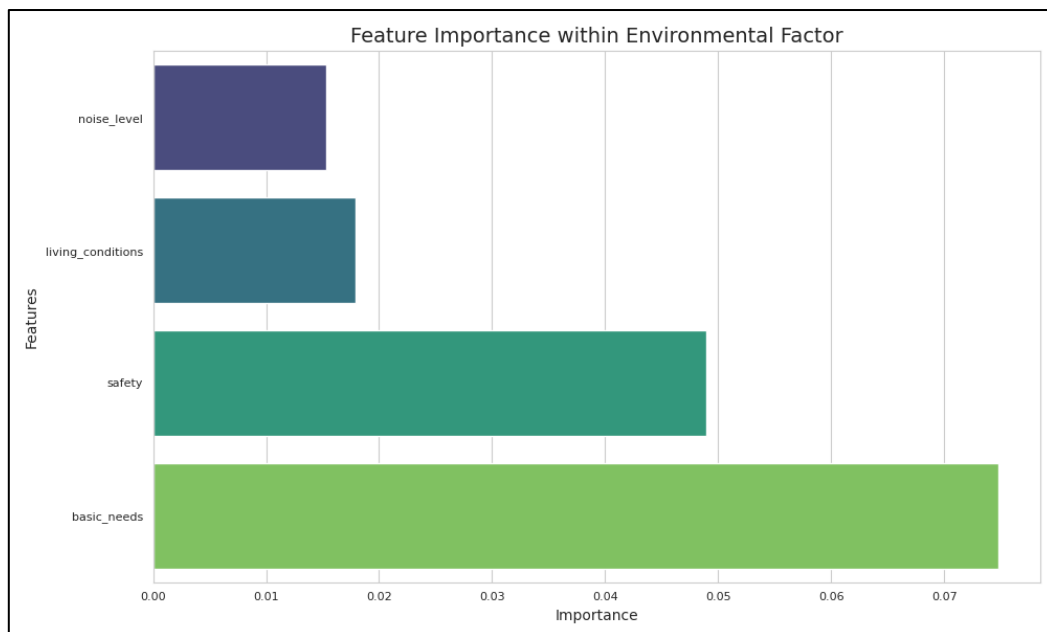


Рис.6. Важливість ознак чинника навколишнього середовища

Аналіз отриманих даних показує, що найбільшу вагу у структурі чинника навколишнього середовища мають базові потреби (basic needs). Його показник важливості перевищує 0,7, що вказує на його значний вплив на рівень стресу. Другим за значущістю фактором виявилася безпека (safety), що свідчить про важливість стабільного та захищеного середовища для нормального функціонування особистості, і також досить значний вплив на рівень стресу.

На рисунку 7 зображено важливість ознак психологічних чинників.

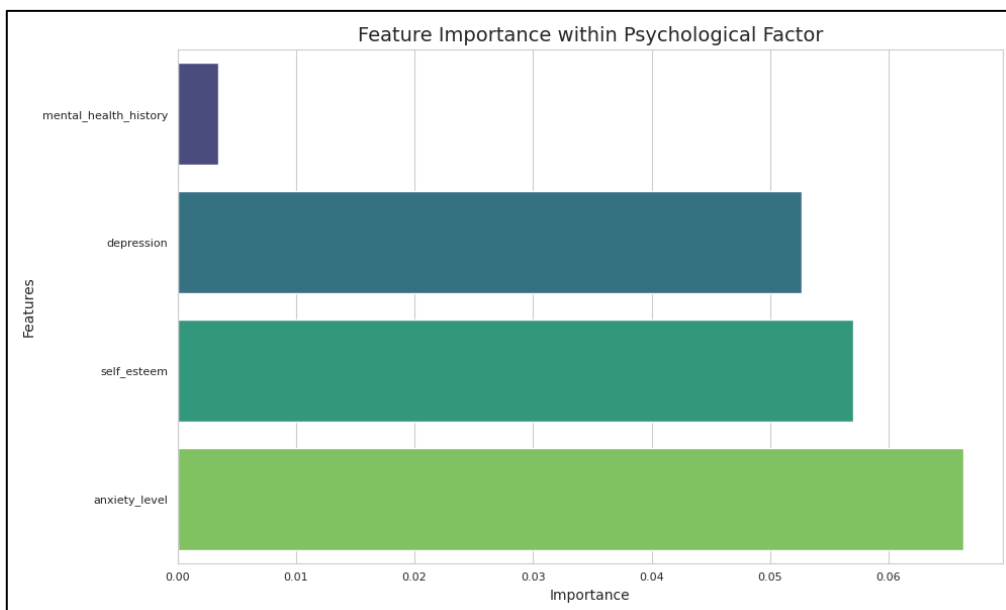


Рис.7. Важливість ознак психологічних чинників

З рисунка 7 бачимо, що рівень тривоги (anxiety level) є найбільш важливою ознакою, маючи найбільше значення, що вказує на її домінуючий вплив на рівень стресу у студентів. Самооцінка (self esteem) є наступним за важливістю психологічним фактором, що суттєво впливає на загальний стан і також має значний вплив на рівень стресу у студентів. Депресія (depression), зі значенням близько 0,057, також є дуже значущою у впливі на рівень стресу у студентів.

Далі візуалізуємо важливість ознак соціальних чинників (рис.8).

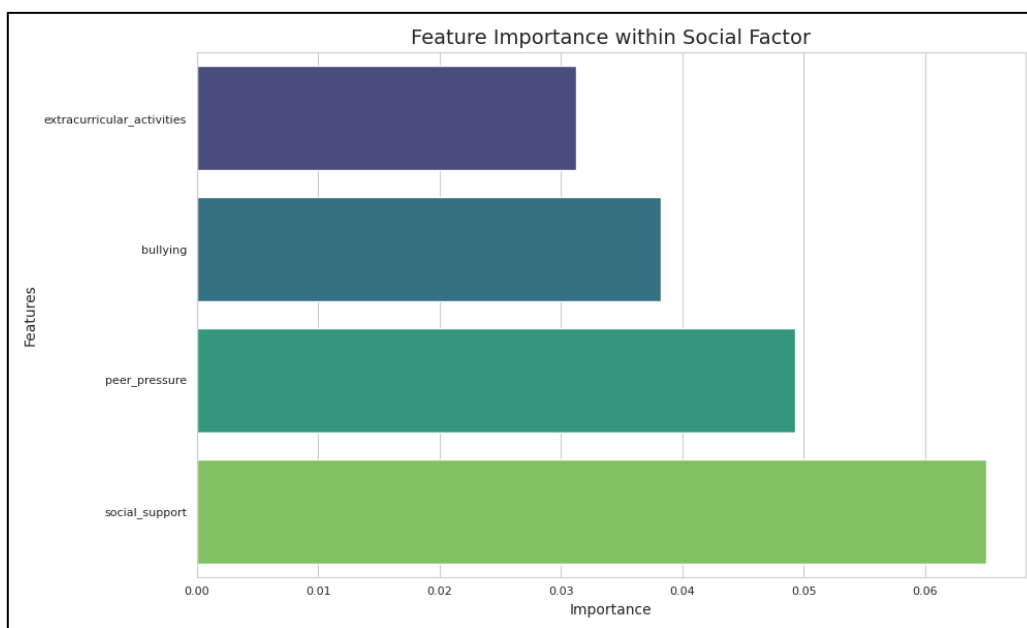


Рис.8. Важливість ознак соціальних факторів

Аналіз діаграми показує, що найбільшу вагу в структурі соціального фактора має соціальна підтримка (social support), що свідчить про те, що вона найбільше впливає на стрес. На другому місці за значущістю перебуває тиск з боку однолітків (peer pressure), що вказує на суттєвий вплив взаємодії з однолітками на ймовірність виникнення стрес. Булінг (bullying) має дещо нижчий рівень важливості, проте залишається значущим фактором. Найменш впливовою ознакою є позакласна активність (extracurricular activities), що може бути наслідком їхнього меншого прямого впливу на стрес.

Цільовою змінною в даному дослідженні є Stress_level, яка відображає рівень стресу студента і представлена у вигляді класів (наприклад, 0 – низький, 1 – середній, 2 – високий). Було детально проаналізовано розподіл студентів за цими рівнями стресу.

На рисунку 9 зображено кругову діаграму, яка відображає кількісний розподіл параметру Stress_level у відсотках.

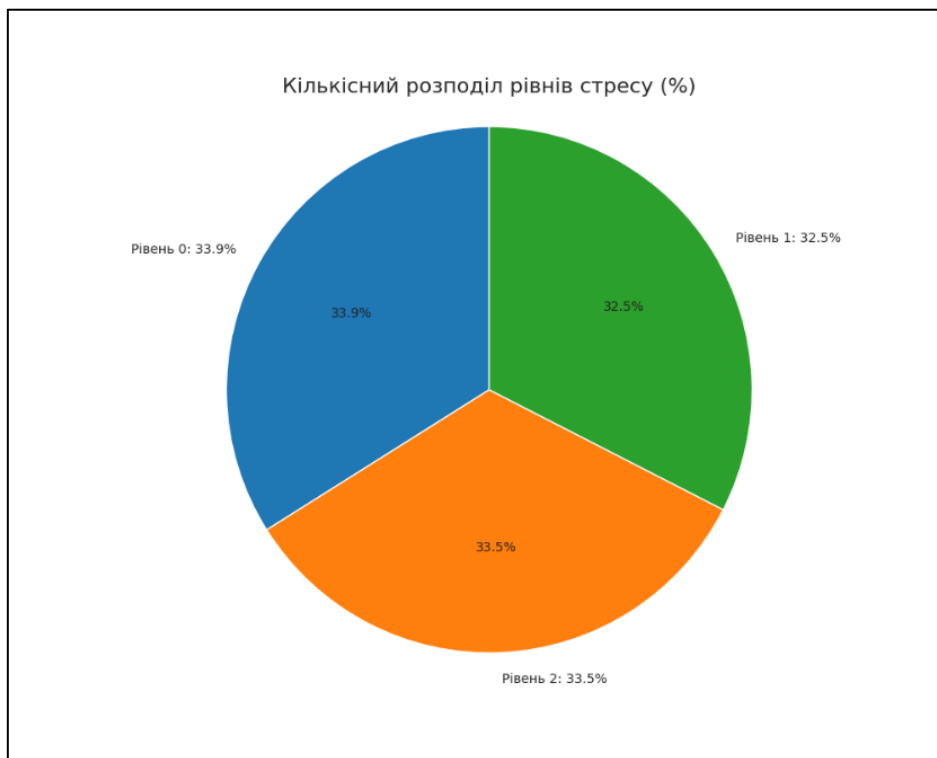


Рис.9. Кількісний розподіл рівнів стресу у відсотках

На основі аналізу кількісних показників рівня стресу можна зробити висновок, що значення цього параметра серед студентів закладів вищої освіти в представленому датасеті мають рівномірний розподіл.

Висновки

У ході розвідувального аналізу даних було досліджено фактори, що впливають на рівень стресу у студентів закладів вищої освіти, а саме психологічні, фізіологічні, соціальні, академічні та фактори навколишнього середовища. Аналіз показав, що найбільш впливовими факторами, що збільшують ймовірність виникнення стресу, є: тривога, депресія, булінг, проблеми зі сном, фінансові труднощі та занепокоєння щодо майбутньої кар'єри. У той час, такі аспекти, як висока самооцінка, хороша академічна успішність та позитивні взаємини з викладачами, понижують рівень стресу.

Побудовано матрицю кореляції, яка показує залежність між усіма параметрами та зроблено висновки щодо залежності параметру рівня стресу до інших ознак. Було побудовано стовпчасті діаграми для візуалізації кількості даних відповідно до окремих ознак, а саме рівень тривоги, самооцінка, депресія, якість сну, навчальне навантаження, майбутня кар'єра, рівень соціальної підтримки та ін.

Також виконано візуалізацію кількості даних відповідно до рівня стресу за допомогою кругової діаграми, що показало нам рівномірне розподілення даних.

Таким чином, вплив стресу на студентів є комплексним і багатовимірним. Це вимагає впровадження системних рішень – зокрема, моніторингу рівня стресу, профілактики емоційного вигорання, формування навичок стресостійкості, а також використання інноваційних цифрових інструментів, зокрема методів машинного навчання, для своєчасного виявлення та підтримки студентів груп ризику.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Хоменко-Семенова Л. О., Прохоренко Я. С. Вплив інформаційного контенту на емоційну стійкість студентів ЗВО в умовах воєнного стану. *Вісник Національного авіаційного університету*. Серія: Педагогіка, Психологія, 2022, Випуск 1 (20). С. 129–134. URL: <https://er.nau.edu.ua/server/api/core/bitstreams/6967222c-b23a-402c-b7e3-101e5b7bb698/content>
2. Нагірнер І. І. Умови формування психологічної стійкості в навчальному процесі ВНЗ. *Вісник Львівського університету. Серія психологічні науки*. 2022. Випуск 14. С. 3–14. DOI: <https://doi.org/10.30970/PS.2022.14.1>
3. Böke B. N., Mills D. J., Mettler J., Heath N. L. Stress and Coping Patterns of University Students, *Journal of College Student Development*. 2019. 60(1). P. 85-103. URL: <https://dx.doi.org/10.1353/csd.2019.0005>
4. Gong Z, Wang H, Zhong M, Shao Y. College students' learning stress, psychological resilience and learning burnout: status quo and coping strategies. *BMC Psychiatry*. 2023. Article 389. DOI: [10.1186/s12888-023-04783-z](https://doi.org/10.1186/s12888-023-04783-z)
5. Nguyen T., Pham, H., Le, V. Family and Academic Stress and Their Impact on Student's Depression and Academic Performance. *Journal of Education and Health Promotion*. 2022. Article 16:13:869337. DOI: [10.3389/fpsy.2022.869337](https://doi.org/10.3389/fpsy.2022.869337)
6. Chhabii. Student Stress Factors: A Comprehensive Analysis. 2024. URL: <https://www.kaggle.com/datasets/rxnach/student-stress-factors-a-comprehensive-analysis>
7. Мокін В. Б., Дратовий М. В. Наука про дані: машинне навчання та інтелектуальний аналіз даних. Навчальний посібник. м. Вінниця. 2024. 263 с. URL: <https://iq.vntu.edu.ua/repository/getfile.php/8163.pdf>
8. A Complete Guide to Data Visualization in Python With Libraries & More. 2025 URL: <https://www.simplilearn.com/tutorials/python-tutorial/data-visualization-in-python>
9. Seaborn Tutorial. 2025. URL: <https://seaborn.pydata.org/tutorial.html>
10. Matplotlib Pyplot Documentation. 2025. URL: https://matplotlib.org/3.5.3/api/as_gen/matplotlib.pyplot.html

Верчинська Марина Олександрівна – студентка групи 2ICT-216, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: marinkdverchynska@gmail.com

Войцеховська Ольга Олександрівна – PhD, старший викладач кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, м. Вінниця, e-mail: olgav1085@gmail.com

Verchynska Maryna O. – student of group 2IST-2b, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: marinkdverchynska@gmail.com

Voitsekhovska Olha O. – PhD, Senior Lecturer of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: olgav1085@gmail.com