

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДВНЗ «ПРИКАРПАТСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАСИЛЯ СТЕФАНІКА»**

Факультет математики та інформатики

Кафедра інформаційних технологій

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Основи науки про дані

Освітня програма «Якість та безпека програмного забезпечення»

Спеціальність 121 Інженерія програмного забезпечення

Галузь знань 12 Інформаційні технології

Затверджено на засіданні кафедри
Протокол № 1 від 31.08.2020 р.

м. Івано-Франківськ - 2020

ЗМІСТ

1	Загальна інформація.....	3
2	Анотація до курсу.....	3
3	Мета та цілі курсу.....	3
4	Компетентності.....	3
5	Результати навчання.....	3
6	Організація навчання курсу.....	4
7	Система оцінювання курсу.....	5
8	Політика курсу.....	5
9	Рекомендована література.....	6

1. Загальна інформація

Назва дисципліни	Основи науки про дані
Рівень вищої освіти	Другий (магістерський)
Викладач (-і)	Козленко Микола Іванович, доцент кафедри інформаційних технологій, канд. техн. наук, доцент
Контактний телефон викладача	+380 (342) 59-60-58
Е-mail викладача	mykola.kozlenko@pnu.edu.ua
Формат дисципліни	дисципліна вільного вибору студента
Обсяг дисципліни	6 кредитів ECTS
Посилання на сайт дистанційного навчання	https://ceeq.pnu.edu.ua
Консультації	Вівторок 15.00 год. 319 ауд. адміністративного корпусу

2. Анотація до курсу

Предметом вивчення навчального курсу «Основи науки про дані» є наукові методи, процеси і системи, які стосуються добування знань із даних у різних формах, як структурованих так і неструктурованих. Даний курс базується на курсах «Програмування мовою Пайтон», «Основи штучного інтелекту», «Інтелектуальний аналіз даних» першого (бакалаврського) рівня освіти.

3. Мета та цілі курсу

Метою курсу “Основи науки про дані” є ознайомлення студентів з основними теоретичними положеннями про принципи, методик і методи що базуються на даних. Цілі курсу: вивчення принципів та методів науки про дані стосовно обробки сигналів, ймовірнісних моделей, машинного та статистичне навчання, класифікації, кластеризації, добування даних, баз даних, теорії розпізнавання образів, візуалізації, передбачувальної аналітики, нечіткого моделювання, сховищ даних, стиснення даних, програмування, штучного інтелекту загалом.

4. Компетентності

ЗК-1. Здатність до абстрактного мислення, аналізу та синтезу.

5. Результати навчання

ПР-1. Знати і системно застосовувати методи аналізу та моделювання прикладної області, виявлення інформаційних потреб і збору вихідних даних для проектування програмного забезпечення.

ПР-6. Аналізувати, оцінювати і вибирати методи, сучасні програмно-апаратні інструментальні та обчислювальні засоби, технології, алгоритмічні та програмні рішення для ефективного виконання конкретних виробничих задач з програмної інженерії.

ПР-7. Обґрунтовано вибирати парадигми і мови програмування для вирішення прикладних завдань; застосовувати на практиці системні та спеціалізовані засоби, компонентні технології (платформи) та інтегровані середовища розробки програмного забезпечення.

ПР-8. Проводити аналітичне дослідження параметрів функціонування програмних систем для їх валідації та верифікації, а також проводити аналіз обраних методів, засобів автоматизованого проектування та реалізації програмного забезпечення.

6. Організація навчання курсу

Обсяг курсу

Вид заняття	Загальна кількість годин
лекції	30
лабораторні	30
самостійна робота	120

Ознаки курсу

Семестр	Спеціальність	Рік навчання	Нормативний / вибірковий
3	121	2	вибірковий

Тематика курсу

Тема, план	Форма заняття	Література	Завдання, год	Вага оцінки	Термін виконання
1. Основні поняття теорії ймовірностей	Лекція	[1]	2	0,02	1 тиждень
	Лаб.		2	0,03	
2. Основні поняття та методи мат. статистики	Лекція	[2]	2	0,02	1 тиждень
	Лаб.		2	0,03	
3. Data acquisition, data entry, signal reception, data extraction	Лекція	[3]	2	0,02	1 тиждень
	Лаб.		2	0,03	
4. Data warehousing, очистка даних	Лекція	[4]	2	0,02	1 тиждень
	Лаб.		2	0,03	
5. Data mining, кластеризація, класифікація	Лекція	[5]	2	0,02	1 тиждень
	Лаб.		2	0,03	
6. Презентація та візуалізація даних	Лекція	[6]	2	0,02	1 тиждень
	Лаб.		2	0,03	
7. Розвідувальний аналіз даних, передбачувальна аналітика	Лекція	[7]	2	0,02	1 тиждень
	Лаб.		2	0,03	
8. Статистичне навчання	Лекція	[8]	2	0,02	1 тиждень
	Лаб.		2	0,03	
9. Машинне навчання	Лекція	[9]	2	0,02	1 тиждень
	Лаб.		2	0,03	
10. Глибинне навчання	Лекція	[10]	2	0,02	1 тиждень
	Лаб.		2	0,03	
11. Виявлення аномалій	Лекція	[11]	2	0,02	1 тиждень
	Лаб.		2	0,03	
12. Передбачення	Лекція	[12]	2	0,02	1 тиждень
	Лаб.		2	0,03	
13. Виявлення патернів	Лекція	[13]	2	0,02	1 тиждень
	Лаб.		2	0,03	
14. Розпізнавання сутностей	Лекція	[14]	2	0,02	1 тиждень
	Лаб.		2	0,03	
15. Рекомендаційні системи	Лекція	[15]	2	0,02	1 тиждень
	Лаб.		2	0,03	

7. Система оцінювання курсу

Загальна система оцінювання курсу	Сума балів	Оцінка ECTS	Оцінка за національною шкалою
	90 – 100	A	зараховано
	80 – 89	B	зараховано
	70 – 79	C	
	60 – 69	D	зараховано
	50 – 59	E	
	26 – 49	FX	незараховано з можливістю повторного складання
	0 – 25	F	незараховано з обов'язковим повторним вивченням дисципліни
Вимоги до письмових робіт	Контрольні робота в тестовій формі за кожною темою (30 балів)		
Лабораторні роботи	Лабораторні роботи за кожною темою (45 балів)		
Самостійна робота	Сертифікат про неформальну освіту або науково-дослідна робота або індивідуальне завдання (25 балів)		

8. Політика курсу

Студент, перед виконанням лабораторних робіт, отримує індивідуальне завдання та самостійно працює над його виконанням. За результатами виконання лабораторної роботи здається звіт, який захищається усно. Це сприяє розвитку навичок самостійної роботи над поставленою задачею та індивідуальному підходу у опануванні курсу із врахуванням можливостей та базового рівня студента.

Академічна доброчесність. Дотримання академічної доброчесності студентами передбачає:

- самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю результатів навчання;
- посилення на джерела інформації у разі використання ідей, розробок, тверджень, відомостей;
- надання достовірної інформації про результати власної навчальної (наукової, творчої) діяльності, використанні методики досліджень і джерела інформації.

Порушенням академічної доброчесності вважається:

- академічний плагіат оприлюднення (частково або повністю) наукових (творчих) результатів, отриманих іншими особами, як результатів власного дослідження (творчості) та/або відтворення опублікованих текстів (оприлюднених творів мистецтва) інших авторів без зазначення авторства;
- самоплагіат оприлюднення (частково або повністю) власних раніше опублікованих наукових результатів як нових наукових результатів; фабрикація - вигадкування даних чи фактів, що використовуються в освітньому процесі або наукових дослідженнях;
- фальсифікація свідомо зміна чи модифікація вже наявних даних, що стосуються освітнього процесу чи наукових досліджень;
- списування виконання письмових робіт із залученням зовнішніх джерел інформації, крім дозволених для використання, зокрема під час оцінювання результатів навчання.

За порушення академічної доброчесності здобувачі освіти можуть бути притягнені до такої академічної відповідальності: повторне проходження оцінювання (контрольна робота, іспит тощо); повторне проходження відповідного освітнього компонента освітньої програми.

9. Рекомендована література

1. Grinstead, C.M. and Snell, J.L., 2012. *Introduction to probability*. American Mathematical Soc.
2. Hogg, R.V., McKean, J. and Craig, A.T., 2005. *Introduction to mathematical statistics*. Pearson Education.
3. Di Paolo Emilio, M., 2013. *Data acquisition systems: from fundamentals to applied design*.
4. Devlin, B. and Cote, L.D., 1996. *Data warehouse: from architecture to implementation*. Addison-Wesley Longman Publishing Co., Inc.
5. Kantardzic, M., 2011. *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons.
6. Post, F.H., Nielson, G. and Bonneau, G.P. eds., 2002. *Data visualization: The state of the art* (Vol. 713). Springer Science & Business Media.
7. Cox, V., 2017. Exploratory data analysis. In *Translating Statistics to Make Decisions* (pp. 47-74). Apress, Berkeley, CA.
8. James, G., Witten, D., Hastie, T. and Tibshirani, R., 2013. *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.
9. Hackeling, G., 2017. *Mastering Machine Learning with scikit-learn*. Packt Publishing Ltd.
10. Remember, A.N.T. and Asians, C.R., *Machine Learning For Hackers Case Studies And Algorithms To Get You Started*.
11. Dunning, T. and Friedman, E., 2014. *Practical machine learning: a new look at anomaly detection*. " O'Reilly Media, Inc."
12. Brockwell, P.J., Davis, R.A. and Calder, M.V., 2002. *Introduction to time series and forecasting* (Vol. 2, pp. 3118-3121). New York: springer.
13. Ripley, B.D., 2007. *Pattern recognition and neural networks*. Cambridge university press.
14. Bunke, H. and Sanfeliu, A., 1990. *Syntactic and structural pattern recognition: theory and applications* (Vol. 7). World Scientific.
15. Manouselis, N., Drachsler, H., Verbert, K. and Duval, E., 2012. *Recommender systems for learning*. Springer Science & Business Media.

Посилання на результати власних наукових досліджень викладачів за змістом курсу

1. M. Kozlenko, I. Lazarovych, V. Tkachuk and V. Vialkova, "Software Demodulation of Weak Radio Signals using Convolutional Neural Network," *2020 IEEE 7th International Conference on Energy Smart Systems (ESS)*, Kyiv, Ukraine, 2020, pp. 339-342, doi: 10.1109/ESS50319.2020.9160035.
2. M. Kozlenko, V. Tkachuk, and M. Dutchak, "Software implementation of microcomputer based intrusion detection and prevention system with binary neural network," in *Proceedings 2nd International Scientific-Practical Conference on Problems of Cyber Security of Information and Telecommunication Systems (PCSITS)*, O. Oksiiuk et al, Eds. Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, Apr. 11-12, 2019, pp. 371-373.
3. M. Kozlenko, I. Lazarovych, and M. Kuz, "Named entity recognition software for automated extraction of call signs and service messages in shortwave radio communications," in *Proceedings of the 3rd International Scientific-Practical Conference on Problems of Cyber Security of Information and Telecommunication Systems (PCSITS)*, O. Oksiiuk et al, Eds. Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, June 12, 2020, pp. 72-73.
4. M. Kozlenko and V. Tkachuk, "Deep learning based detection of DNS spoofing attack," in *Proceedings of the 2019 Scientific Seminar on Innovative Solutions in Software Engineering*, Ivano-Frankivsk, Ukraine, Dec. 10, 2019, pp. 10-11.

5. M. Kozlenko, A. Bosyi, O. Simkiv, and N. Savchenko, "Artificial intelligence in e-commerce," in *Proc. 3rd International Conference on Applied Information Systems and Technologies in the Information Society (AISTIS)*, V. Pleskach and V. Mironova, Eds. Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, Sept. 30, 2019, pp. 207-211.
6. M. Kozlenko, I. Lazarovych, and M. Kuz, "Deep learning approach to signal processing in infocommunications," in *Proc. 4th International Scientific and Practical Conference on Applied Systems and Technologies in the Information Society (AISTIS)*, V. Pleskach and V. Mironova, Eds. Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, Sept. 30, 2020, pp. 81-82.

Наукові роботи провідних науковців світового рівня

1. S. Kumar, N. Dhanda and A. Pandey, "Data Science — Cosmic Infoset Mining, Modeling and Visualization," *2018 International Conference on Computational and Characterization Techniques in Engineering & Sciences (CCTES)*, Lucknow, India, 2018, pp. 1-4, doi: 10.1109/CCTES.2018.8674138.
2. C. Silva, M. Saraee and M. Saraee, "Data Science in Public Mental Health: A New Analytic Framework," *2019 IEEE Symposium on Computers and Communications (ISCC)*, Barcelona, Spain, 2019, pp. 1123-1128, doi: 10.1109/ISCC47284.2019.8969723.
3. D. E. Drummond, "Open sourcing education for Data Engineering and Data Science," *2016 IEEE Frontiers in Education Conference (FIE)*, Erie, PA, USA, 2016, pp. 1-1, doi: 10.1109/FIE.2016.7757517.
4. I. B. Hassan and J. Liu, "Embedding Data Science into Computer Science Education," *2019 IEEE International Conference on Electro Information Technology (EIT)*, Brookings, SD, USA, 2019, pp. 367-372, doi: 10.1109/EIT.2019.8833753.
5. M. Muniswamaiah, T. Agerwala and C. C. Tappert, "Federated Query processing for Big Data in Data Science," *2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA, 2019, pp. 6145-6147, doi: 10.1109/BigData47090.2019.9005530.
6. L. Getoor, "Responsible Data Science," *2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA, 2019, pp. 1-1, doi: 10.1109/BigData47090.2019.9006129.
7. A. D. C. Gonzales, F. L. P. Villantoy and D. S. M. Sanchez, "Data Science Model for the Evaluation of Customers of Rural Savings Banks without Credit History," *2019 7th International Engineering, Sciences and Technology Conference (IESTEC)*, Panama, Panama, 2019, pp. 329-334, doi: 10.1109/IESTEC46403.2019.00067.
8. D. Stahl and D. Stamate, "Data Science Challenges in Computational Psychiatry and Psychiatric Research," *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, Turin, Italy, 2018, pp. 524-525, doi: 10.1109/DSAA.2018.00067.
9. C. Phethean, E. Simperl, T. Tiropanis, R. Tinati and W. Hall, "The Role of Data Science in Web Science," in *IEEE Intelligent Systems*, vol. 31, no. 3, pp. 102-107, May-June 2016, doi: 10.1109/MIS.2016.54.
10. Q. Cheng, F. Lopez and A. Hadjixenofontos, "Integrating Introductory Data Science into Computer and Information Literacy through Collaborative Project-based Learning," *2019 IEEE Frontiers in Education Conference (FIE)*, Covington, KY, USA, 2019, pp. 1-5, doi: 10.1109/FIE43999.2019.9028683.

Викладач _____

