



MINISTRY
OF EDUCATION AND SCIENCE
OF UKRAINE



ACADEMY

OF TECHNICAL SCIENCES OF UKRAINE

CONNECTIVE
technologies

DEPARTMENT OF

*INFORMATION
TECHNOLOGY*

2022 International Conference on Innovative Solutions in Software Engineering (ICISSE)

Conference Proceedings

November 29-30, 2022
Ivano-Frankivsk, Ukraine

*Issued by the decision of the Academic Board of the Faculty of Mathematics and Computer Science
(protocol #15/2022)*

This book and its papers are published under the Creative Commons License Attribution 4.0 International
(CC BY 4.0)

Editors

Mykola Kuz  and **Mykola Kozlenko** 

Department of Information Technology
Faculty of Mathematics and Computer Science
Vasyl Stefanyk Precarpathian National University
57 Shevchenka Street
Administration Building, Office #319
Ivano-Frankivsk, 76018, Ukraine
Tel: +380 (342) 59-60-58
E-mail: icisse@pnu.edu.ua

Papers are presented in authors' edition. The authors are responsible for the content of their paper. The organizers, committees, reviewers, or editors reserve the right not to be responsible for the topicality, correctness, completeness or quality of the information provided.

2022 International Conference on Innovative Solutions in Software Engineering (ICISSE), Mykola Kuz and Mykola Kozlenko Eds., Ivano-Frankivsk, Ukraine: Vasyl Stefanyk Precarpathian National University, 2022, 317 p. [E-book]

ISBN 978-966-640-534-3

DOI 10.5281/zenodo.7502536

The high-quality refereed papers appearing in this book compose the proceedings of the 2022 International Conference on Innovative Solutions in Software Engineering (ICISSE). ICISSE is the forum organized by the Department of Information Technology of the Vasyl Stefanyk Precarpathian National University. ICISSE-2022 is held from November 29 to 30, 2022 in Ivano-Frankivsk, Ukraine. The event is intended to bring together researchers, scientists, and engineers to discuss experimental and theoretical results, mainly in the area of software engineering. The conference also covers topics related to computer science, computer engineering, systems analysis, cybersecurity, information systems and technology, industrial automation, electronics, metrology, micro and nanosystems, telecommunications, radio frequency engineering, IT entrepreneurship, and IT education. The current edition of the event is dedicated to the 50th birth anniversary of famous Ukrainian scientist working in information technology, Founder and first Head of the Department of Information Technology, Vice-Rector for Research of the university, Professor Dr. Pavlo Fedoruk, who passed away in 2013. The conference is included into the list of scientific conferences scheduled for 2022 by the Ministry of Education and Science of Ukraine.

Recommended paper citation template:

J. K. Author, "Decapitalized title of the paper," *2022 International Conference on Innovative Solutions in Software Engineering (ICISSE)*, Vasyl Stefanyk Precarpathian National University, Ivano-Frankivsk, Ukraine, Nov. 29-30, 2022, pp. xxx-xxx.

ISBN 978-966-640-534-3

© Copyright for the individual papers by the papers' authors, 2022
© Copyright for the book as a collection by its editors, 2022
© Vasyl Stefanyk Precarpathian National University, 2022

COMMITTEES

General Chair:

Prof. Dr. Mykola Kuz

Technical Program Committee Chair:

Assoc. Prof. PhD Mykola Kozlenko

Local Organizing Committee Chair:

Mariia Dutchak

Publicication Chair:

Assoc. Prof. PhD Mykola Kozlenko

Preface

The 2022 International Conference on Innovative Solutions in Software Engineering dedicated to the 50th anniversary of an outstanding researcher and teacher, the founder of the Department of Information Technology at Vasyl Stefanyk Precarpathian University, Prof. Dr. Pavlo Fedoruk was held on November 29-30, 2022 in Ivano-Frankivsk, Ukraine.

Vasyl Stefanyk Precarpathian National University has more than 80 years of history. Its inspirational success story was written by people who devoted themselves to science, education, and culture. Among the prominent figures of the University, the name of the outstanding modern Ukrainian scientist Pavlo Fedoruk shone as a bright star. Now he would have celebrated his 50th birthday.

Pavlo Fedoruk graduated from Precarpathian University in 1994. Immediately after graduating, he was offered the position at the Information Department of the University. He proved himself to be a professional in the field of information technology, a talented person with progressive thinking. He was offered the position as a head of the Information Department in October 1994. Since 2002, Pavlo Fedoruk worked as a director of the Center for Information Technology. He also founded the Center of Distance Learning. While performing his official duties, he was simultaneously involved in research. He was a hard worker and independent thinking researcher who worked effectively on telecommunication technology research.

He studied as a PhD student in the period from 1999 to 2002. He received his PhD from the Institute of Mathematical Machines and Systems of the National Academy of Sciences of Ukraine. The topic of the PhD thesis was related to the systems of distance learning and knowledge assessment based on Internet technology. During his work at the University, Pavlo Fedoruk published more than 120 research papers and two monographs. He developed a course on Informatics for students of the Faculty of Economics and published the textbook "Computer for Economists". Also he developed courses on artificial intelligence and intelligent systems for students of the Faculty of Mathematics and Computer Science. In the period from 2006 to 2008, he was a scholarship holder of the Cabinet of Ministers of Ukraine for young scientists. In 2010, he was successfully awarded with the degree of Doctor of Technical Science. The topic of the thesis was related to adaptive systems of distance learning.

Since 2007, he was a member of the program committee and a session chairman at the international scientific and practical conferences "Education based on web technologies" and "Advanced computer technology in education". In 2007, he participated in the International Referee Committee as a representative from Ukraine. In 2008, he was an expert of the Scientific and Technical Council of the Ministry of Education and Science of Ukraine at the Informatics and Cybernetics section. He was a member of specialized academic councils at the Institute of Mathematical Machines and Systems and Ternopil National Economic University. He supervised four research projects in the field of distance learning. In 2009, he was awarded the Tersenov Regional Prize for Innovation from the Society of Inventors and Innovators of Ukraine. In 2010, he was recognized as a Famous Scientist of the Year for significant achievements in the field of research and education and for personal contribution. In 2011, he was awarded an honorary certificate from the Cabinet of Ministers of Ukraine for his significant personal contribution to science.

In 2012, he was promoted to the position of vice-rector for research of the University. His views were based on a broad knowledge and interests. He was a founder of many scientific, social, and entrepreneurial projects. He was one of the founders of the "Firtka" Local News Agency. He was also the President of the Ivano-Frankivsk Regional Martial Arts Federation.

The authors of this collection devoted their papers to the memory of Prof. Dr. Pavlo Fedoruk.

TABLE OF CONTENTS

Paper ID	Paper	Page
2	Moulfi Bouchra, Ferouani Souheyla, Ziani Kerarti Djalal, Marouf Fatima Zohra. New design of gold nano printed antenna with Gallium substrate for optical transmission	1
3	Zeba Shamsi, Subia Shamsi, Sekar K. A survey of the dark web	5
7	Наталія Полукетова, Анатолій Переверзев, Катерина Дереза. Моделювання ефективності дистанційної освіти із застосуванням агентного підходу	7
8	Oleh Zaritskyi. A practical approach to evaluating the ethics index of artificial intelligence systems	15
9	Богдан Савенко. Розподілені системи виявлення зловмисного програмного забезпечення	22
10	Олександр Мельниченко. Метод керування та самоорганізована система планування маршрутів БПЛА	26
11	Ольга Дмитрієва, Ілля Бутенко. Розробка веб-додатку для дистанційного навчання	31
12	Андрій Яворський, Максим Карпаш, Віталій Цих, Любомир Жовтуля. Пілотна система енергетичного моніторингу для навчальної будівлі університету	34
13	Iryna Tarasevich, Alena Tarasevich, Mykhailo Kuzmenko. Impact of Covid-19 on university education digitalization: a critical interpretive synthesis of tensions and barriers for automation field	43
14	Olga Solovei. Selection of data preprocessing technique for imbalanced dataset	48
15	Катерина Островська, Іван Стівченко, Сергій Черський. Дослідження можливості використання нейромережових технологій для вирішення задачі класифікації дефектів металопрокату.	51
16	Катерина Островська, Денис Печений. Дослідження методів на основі нейронних мереж для аналізу тональності корпусу текстів	55
17	Катерина Островська, Анастасія Петренко. Дослідження ефективності методів обробки зображень у системах розпізнання облич	59
18	Анастасія Нічепорук, Андрій Нічепорук, Андрій Казанцев, Олег Парніцький. MitM кібератаки на локально-керуючий рівень кіберфізичної системи розумних мереж енергозабезпечення: визначення доменної області та поведінка системи	63
19	Олександр Мельников, Вікторія Денисенко. Постановка задачі розробки програмного забезпечення для оцінювання лісистості та оброблення інформації про лісові насадження	67
20	Viktor Bilishchuk. Application of virtual simulation for device design	71
21	Iryna Vashchyshak, Serhii Vashchyshak. The web application of calculation of parameters for a small hydroelectric power plant	74
22	Olena Hladka, Ivan Karpovich, Daria Buryan. Modeling technologies for assessment of information security risks in e-commerce	78
24	Михайло Музикін, Галина Нестеренко, Олег Стрелко, Ніна Клочкова. Роль навігаційних систем в організації ефективної роботи автомобільного транспорту	82
25	Борис Незамай. Прогнозування якості програмного забезпечення	86
26	Oleg Zhulkovskiy, Inna Zhulkovska, Serhii Panteikov, Kyrylo Muzychka, Olena Panteikova, Yulia Pavlenko. Mathematical model for studying the thermal regime of the slag-splashing lance	90
27	Анастасія Чучвара, Ольга Чернуха, Юрій Білушак. Системний підхід до математичного опису дифузійних процесів у випадково неоднорідних двофазних структурах з кульовими включеннями	93
28	Oleh Pihnastyi, Maksym Sobol, Dmytro Yelchaninov. On the characteristics of the input material flow of the transport conveyor	102
29	Pavlo Katin, Oleksandr Pokhlyenko. Use of state pattern with the ability to dynamic add new states for software part of automatic cleaning-in-place stations for fermentation tanks	107
30	Petro Kostrobij, Vitalii Poloyi, Iryna Ryzha. Behaviour of SPP waves frequency spectrum for dielectric/metal/dielectric structures	115
32	Vadym Malinovskiy, Leonid Kupershtein, Vitaliy Lukichov. Risks assessment and approaches to creative of the reliable software modules for IoT devices	122
33	Михайло Крихівський, Любов Саманів. Математична модель інформаційного забезпечення екологічних служб	128
34	Roman Dorosh. Application of blockchain and the SimHash algorithm to detect Plagiarism	130

35	Василь Олійник, Андрій Подорожняк, Наталія Любченко. Аналіз та розпізнавання спаму в соціальних мережах та месенджерах	135
36	Олександр Жеребцов, Арсеній Рассохін. Проектування виконавчого пристрою системи контролю управління доступу на базі STM32F103	139
37	Юлія Казимиренко, Оксана Гайдаєнко, Владислав Принць, Анастасія Байталюк. Застосування нечітких алгоритмів в задачах розробки інформаційних систем для прогнозування результатів кіберспортивних змагань	147
38	Андрій Павлишин, Володимир Грига, Володимир Мандзюк. Реалізація алгоритму швидкого перетворення Фур'є на ПЛІС	149
39	Oleksandr Tsurkan, Yuliia Prokop. Automatic number plate recognition from videostream	157
40	Olena Oleinik, Ihor Michurin. Using Goroutine to accelerate SQL queries for PostgreSQL in AWS	160
41	Олександр Жеребцов, Антон Іжиков. Використання функціонального програмування у JavaScript та фреймворках	162
43	Денис Гобов. Контекст IT-проекту та проблеми бізнес-аналізу	169
45	Олександр Жеребцов, Марк Бороденко. Створення багатофункціонального локального веб-серверу на базі ESP8266	173
46	Roman Melnyk, Ruslan Tushnytskyu, Andriy Shpek. Selection of defective chains in a PCB image by thinning, flood-filling and overlaying	179
47	Віра Гарасимів. Автоматизоване тестування в Selenium: Page Object та Page Factory	184
48	Олександр Мельников, Дмитро Козуб. Постановка задачі розробки програмного забезпечення для оцінювання ефективності протиепідемічних заходів та прогнозування зміни відсотка інфікованих та перенесених хвороб у тяжкій формі	186
49	Юлія Сурганова, Юрій Міхлін. Нелінійні нормальні моди маятникової системи під дією магнітних збуджень	190
50	Ярослав Марчук. Дослідження використання паралельних технологій у високо-навантажувальних додатках	193
51	Ivan Rzhetskyi, Anastasiya Troynina. Garbage classification for mobile devices based on MobileNet	201
52	Наталія Іванічок, Олег Іванічок, Софія-Вікторія Склепова, Павло Колковський, Богдан Рачій, Роман Лісовський. Моделювання імпедансних спектрів, отриманих для системи пористий електрод/електроліт	205
53	Роман Базилевич, Андрій Франко. Класифікація структурних елементів програмного забезпечення у задачах генерації тестів	209
54	Роман Боднар. Пристрій для дослідження поверхневого натягу рідин методом газового струменя	212
56	Nataliia Liubchenko, Andrii Podorozhniak, Daniil Onishchenko. Development of forestry application using deep learning technology	216
57	Валентин Соболев, Андрій Подорожняк. Алгоритми пошуку шляху для безпілотних сільськогосподарських систем	218
58	Ольга Чорна, Данило Непорада, Олексій Чорний, Денис Калінін. Мобільний додаток «Система моніторингу асинхронних двигунів» на базі ОС Android	220
59	Тетяна Селівьорстова, Михайло Калюжний, Юрій Каліберда. Програма реалізація сервісу агрегації даних для дослідження комерційних ринкових пропозицій автомобільного ринку України	224
60	Сергій Альошин, Олена Гайтан. Інтелектуальна технологія синтезу автоматизованої експертної системи керування підприємством	226
61	Mukhailo Kvasnii, Vladyslav Shevchuk, Halyna Klym. Intelligent system for collecting and analyzing of meteorological data based on robotic platform	232
63	Яна Лебеденко, Юрій Міхлін. Дослідження стаціонарного режиму та перехідного процесу в неідеальних системах	235
66	Вікторія Пахомова. Прогнозування часу доставки даних в ІТС залізничного транспорту з використанням нейронної технології	238
68	Vasyl Franchuk, Nataliia Franchuk. Using Microsoft Family Safety for parents and children	243
69	Олена Смалько. Про сучасне різноманіття технологій та архітектурних рішень Інтернету речей	251
70	Vyacheslav Fomenko, Oksana Shovkopliias. Informational system of relief generation based on Perl's noise mathematical algorithm	254

72	Ihor Polataiko. Software architecture design of software developed in student engineering teams	257
73	Валерій Ткачук, Олександр Кулик. Реалізація еволюційного алгоритму на реальному квантовому комп'ютері	262
75	Марія Дутчак, Микола Пікуляк. Оцінка якості методів та програмних засобів інтелектуальних адаптивних освітніх веб-систем самонавчання	266
76	Микола Пікуляк, Марія Дутчак. Програмні аспекти виявлення академічного плагіату	271
77	Mariia Slobodian, Mykola Kozlenko. Machine learning based animal emotion classification using audio signals	277
78	Yaroslav Drin, Iryna Drin, Roman Drin. The third initial-boundary value problem for the nonlinear equation of heat conduction with deviation of the argument	282
79	Владислав Гой. Дослідження масштабування та розробка рекомендаційної системи методами глибинного навчання та технологій великих даних	286
80	Дмитро Юрчук, Іван Савка. Дослідження шкідливого програмного забезпечення	291
81	Микола Кузь, Ігор Лазарович, Руслан Мішагін. Напрямки удосконалення міжнародних стандартів у сфері оцінки якості програмних продуктів ISO/IEC 9126 та ISO/IEC 14598	294
82	Vasyl Ivanochko. Analysis of methods and existing software implementations for automatic resume generation	298
83	Yuliia Dudii. Overview of methods and existing solutions for task order generation	306
84	Valentyna Yakovyn, Taras Kriukov. Object detection in mobile robotics	310
85	Mykola Kozlenko. Supervised machine learning based signal demodulation in chaotic communications	313

Selection of data preprocessing technique for imbalanced dataset

Olga Solovei

*Department of Information Technology and Applied mathematics
Kyiv National University of Construction and Architecture,
Ukraine*

Abstract—In many domains datasets are imbalance, most often the problem is visible in datasets with a medicine diagnoses' prediction; credit's approval; surveys. The results of machine learning algorithms for imbalanced dataset are biased to majority class, whereas the correct predictions of the smaller class somethings could be most important. The goal of the work is to empirically study which data preprocessing technique: feature selection or oversampling to be applied before to train a classification learning algorithm.

Keywords—*Oversampling, Feature Selection, Area under ROC curve, Random Forest classifier.*

I. INTRODUCTION

In study [1] to improve classification models' results for imbalance datasets are proposed the following data preprocessing strategies: 1) oversampling - the smaller class is oversampled to match the size of the major class; 2) down-sizing – the number of samples in the major class is reduced; 3) learning by recognition - to use unsupervised machine learning algorithms and ignore target class in dataset. From the mentioned three strategies the preference is given to oversampling technique as it makes balance distribution between classes which improves that classifier's performance. However, in study [2] can be found that the most widely used oversampling method - SMOTE is good for a highly dimensional imbalance dataset only when k-nearest neighbors classifier is used and most predictive features are selected.

Because the results of the mentioned studies look controversial more empirical tests are required to understand the optimal approach to preprocessing technique selection.

II. RELATED WORK

In work [3] was proposed to use only feature selection technique for a highly dimensional imbalance gene dataset without comparing its efficiency to the results of the oversampling or down-sizing, therefore the concluded results of the study are not obvious.

To eliminate that gap, in the current work, will be conducted experiments to evaluate the efficiency of data preprocessing techniques: 1) a feature selection; 2) an oversampling by SMOTE; 3) a combination of both.

III. METHODOLOGY

Our datasets are from UCI Machine Learning Repository: "adult", "coli" and "songs" are highly dimensional with imbalance ration between the minority to the majority class: 0.3, 0.06, 0.16 correspondingly. The features' count per each dataset is 81, 86 and 58 and samples' count per each dataset is 48842, 9822, 1409 correspondingly. In the experiments are included the steps: 1) create for each dataset two pairs of the train and test data using different data preprocessing techniques: pairset1 – the dataset is left imbalance but the most predictive features are selected by applying mutual information method [4]; pairset2 - the dataset is balanced by SMOTE method so that a minority class 100% matches the majority and the most predictive features are selected; pairset3 - the dataset is balanced by SMOTE method but all features are included. 2) train Random Forest

2022 International Conference on Innovative Solutions in Software Engineering
Ivano-Frankivsk, Ukraine, November 29-30, 2022

(RF) algorithm and evaluate the prediction quality using metrics: area under ROC (AUC), precision, recall and f1-score on the test subset; 4) compare performance's results based on the metrics' values and ROC curves' graphs.

IV. RESULTS

Random Forest classifier's performance metrics for each dataset are recorded in table 1. ROC convex hull graphs presented on fig 1.

Table 1. RF algorithm's performance results and training time

dataset	Pairset1				Pairset2				Pairset3			
	AUC, %	Precision, %	Recall, %	f1 score, %	AUC, %	Precision, %	Recall, %	f1 score, %	AUC, %	Precision, %	Recall, %	f1 score, %
adult	84	64	51	56	93	84	86	85	93	85	86	85
coli	65	52	51	51	99	96	95	96	99	97	95	96
songs	47	57	54	55	85	81	80	80	90	83	83	83

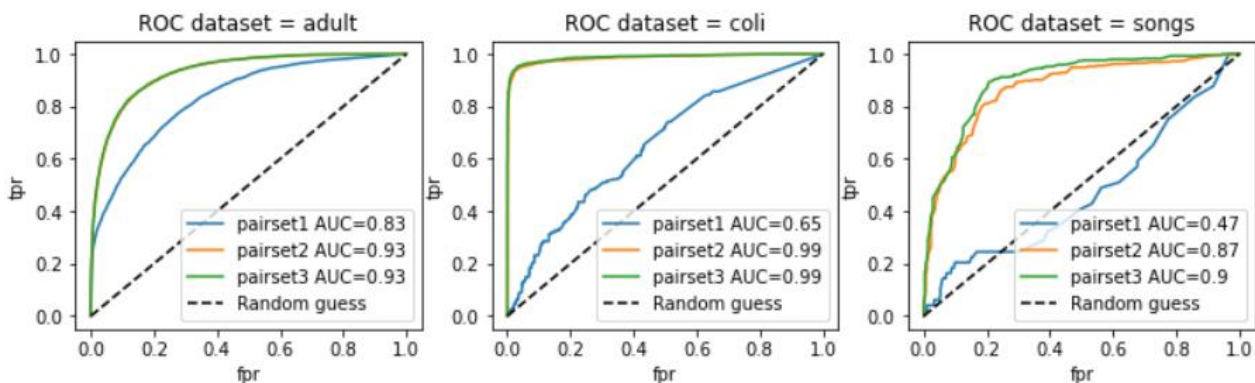


Figure 1. ROC curves of RF classifier trained on pairset1, partset2 and pairset3.

V. DISCUSSION

The results from table 1 can be interpreted as the following: for highly dimensional datasets an ability of RF classifier to predict the target class is better when a dataset is balanced (pairset2, pairset3) compared to the performance of RF on the imbalance dataset with the most predictive features (pairset1). However, for dataset = "coli" the performance of RF on the balance dataset with all features (pairset3) is better compared to the performance of RF on the balance dataset with the most predictive features are selected (pairset2).

The same results are visible on ROCs convex hull graphs from fig 1: the "more northwest" ROC curve is seen for pairset2 and pairset3 on datasets: "adult" and "coli" and the "more northwest" ROC curve is seen for pairset3 on dataset: "songs".

VI. FUTURE RESEARCH

It is worth to further investigate: 1) the patterns in the dataset when applying only feature selection technique can guarantee the high model's performance; 2) how model's performance is affected when down-sizing technique together with feature selection are applied on the data preprocessing stage.

VII. CONCLUSION

Random Forest classifier's performance is better when SMOTE method of the oversampling technique is applied compared to the performance's results received by applying dataset's dimension reduction by selecting the most predictive features using mutual information method.

REFERENCES

- [1] N. Japkowicz. "The class Imbalance Problem: Significance and Strategies," *AAAI conference of Artificial Intelligence*, 2000, pp.10-15.
- [2] R. Blagus, L. Lusa. "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinformatics*, 2013 – Springer, pp.2-16.
- [3] G.A.Sharifai, Z. Zainol. "Feature Selection for High-Dimensional and Imbalanced Biomedical Data Based on Robust Correlation Based Redundancy and Binary Grasshopper Optimization Algorithm," *MDPI Genes*, 2020, pp.-2-26.
- [4] J. Li, K. Cheng, S. Wang, F. Morstatter, R. Trevino, J. Tang & H. Liu. "Feature Selection: A Data Perspective," *ACM Computing Surveys*, vol. (50), no. 6, 2018, pp 1–45, DOI:<https://doi.org/10.1145/3136625>.