Palestinian Journal
of Technology & Applied Sciences (PJTAS)

Vision

Achieving leadership, excellence and innovation in the field of open learning, community service, and scientific research, in addition to reinforcing the University leading role in establishing a Palestinian society built on knowledge and science.

Mission

To prepare qualified graduates equipped with competencies that enable them to address the needs of their community, and compete in both local and regional labor markets. Furthermore, The University seeks to promote students' innovative contributions in scientific research and human and technical capacity-building, through providing them with educational and training programs in accordance with the best practices of open and blended learning approach, as well as through fostering an educational environment that promotes scientific research in accordance with the latest standards of quality and excellence. The University strives to implement its mission within a framework of knowledge exchange and cooperation with the community institutions and experts.

Core Values

To achieve the University's vision, mission and goals, the University strives to practice and promote the following core values:

- Leadership and excellence.
- Patriotism and nationalism.
- Democracy in education and equal opportunities.
- Academic and intellectual freedom.
- Commitment to regulations and bylaws.
- Partnership with the community
- Participative management.
- Enforcing the pioneer role of women.
- Integrity and Transparency.
- Competitiveness.

The Journal

The Palestinian Journal of Technology and Applied Sciences is an annual scientific refereed journal, issued by the Deanship of Graduate Studies and Scientific Research. The first issue of the Journal was published in January 2018 after obtaining an International Standard Serial Number (E- ISSN: 2521-411X), (P– ISSN: 2520-7431).

The journal publishes original research papers and studies conducted by researchers and faculty staff at QOU and by their counterparts at local and overseas universities, in accordance with their academic specializations. The Journal also publishes reviews, scientific reports and translated research papers, provided that these papers have not been published in any conference book or in any other journal.

The Journal comprises the following topics:

Information and Communication Technology, Physics, Chemistry, Biology, Mathematics, Statistics, Biotechnology, Bioinformatics, Agriculture Sciences, Geology, Ecology, Nanotechnology , Mechatronics, Internet of things , Artificial Intelligence and Big Data.

Publication and Documentation Guidelines

First: Requirements of preparing the research:

The research must include the following:

1. A cover page which should include the title of the research stated in English and Arabic, including the name of researcher/researchers, his/her title, and email.
2. Two abstracts (English and Arabic) around (150-200 word). The abstract should include no more than 6 key words.
3. Graphs and diagrams should be placed within the text, serially numbered, and their titles, comments or remarks should be placed underneath.
4. Tables should be placed within the text, serially numbered and titles should be written above the tables, whereas comments or any remarks should be written underneath the tables

Second: Submission Guidelines:

1. The Researcher should submit a letter addressing the Head of Editorial Board in which he/she requests his paper to be published in the Journal, specifying the specialization of his/her paper.
2. The researcher should submit his research via email to the Deanship of Scientific research (tas@qou.edu) in Microsoft Word Format, taking into Consideration that the page layout should be two columns.

   (Check the attached digital form on the website of the Journal)
3. The researcher should submit a written pledge that the paper has not been published nor submitted for publishing in any other periodical, and that it is not a chapter or a part of a published book.
4. The researcher should submit a short Curriculum Vitae (CV) in which she/he includes full name, workplace, academic rank, specific specialization and contact information (phone and mobile number, and e-mail address).
5. Complete copy of the data collection tools (questionnaire or other) if not included in the paper itself or the Annexes.
6. No indication shall be given regarding the name or the identity of the researcher in the research paper, in order to ensure the confidentiality of the arbitration process.

Third- Publication Guidelines:

The editorial board of the journal stresses the importance of the full compliance with the publication guidelines, taking into note that research papers that do not meet the guidelines will not be considered, and they will be returned to the researchers for modification to comply with the publication guidelines.

1. Papers are accepted in English only, and the language used should be well constructed and sound.
2. The researcher must submit his/her research via email (tas@qou.edu)in Microsoft Word format, taking into consideration the following:
   - Font type should be Times New Roman, and the researcher should use bold font size 14 for head titles, bold font size 13 for subtitles, font size 12 for the rest of the text, and font size 11 for tables and diagrams.
   - the text should be single-spaced
   - Margins: Should be set to: 2cm top, 2.5 cm bottom, 1.5 cm left and right.
3. The paper should not exceed 25 (A4) pages or (7000) words including figures and graphics, tables, endnotes, and references, while annexes are inserted after the list of references, though annexes are not published but rather inserted only for the purpose of arbitration.
4. The research has to be characterized by originality, neutrality, and scientific value.
5. The research should not be published or submitted to be published in other journals, and the researcher has to submit a written acknowledgment that the research has never been published or sent for publication in other journals during the completion of the arbitration process. In addition, the main researcher must acknowledge that he/she had read the publication guidelines and he/she is fully abided by them.
6. The research should not be a chapter or part of an already published book.
7. Neither the research nor part of it should be published elsewhere, unless the researcher obtains a written acknowledgement from the Deanship of Scientific Research.
8. The Journal preserves the right to request the researcher to omit, delete, or rephrase any part of his/her paper to suit the publication policy. The Journal has also the right to make any changes on the form/ design of the research.
9. The research must include two research abstracts, one in Arabic and another in English of (150-200) words. The abstract must underline the objectives of the paper, statement of the problem, methodology, and the main conclusions. The researcher is also to provide no more than six keywords at the end of the abstract which enable an easy access in the database.

11. The researcher has to indicate if his research is part of a master thesis or a doctoral dissertation as he/she should clarify this in the cover page, possibly inserted in the footnote.
12. The research papers submitted to the Deanship of Scientific Research will not be returned to the researchers whether accepted or declined.
13. In case the research does not comply with the publication guidelines, the deanship will send a declining letter to the researcher.
14. Researchers must commit to pay the expenses of the arbitration process, in case of withdrawal during the final evaluation process and publication procedures.
15. The researchers will be notified of the results and final decision of the editorial board within a period ranging from three to six months starting from the date of submitting the research.

Four- Documentation:

1. Footnotes should be written at the end of the paper as follows; if the reference is a book, it is cited in the following order, name of the author, title of the book or paper, name of the translator if any or reviser, place of publication, publisher, edition, year of publishing, volume, and page number. If the reference is a journal, it should be cited as follows, author, paper title, journal title, journal volume, date of publication and page number.
2. References and resources should be arranged at the end of the paper in accordance to the alphabetical order starting with the surname of author, followed by the name of the author, title of the book or paper, place of publishing, edition, year of publication, and volume. The list should not include any reference which is not mentioned in the body of the paper.

   - In case the resource is with no specified edition, the researcher writes (N.A)
   - In case the publishing company is in not available, the researcher writes (N.P)
   - In case there is no author, the researcher writes (N.A)
   - In case the publishing date is missing , the researcher writes (N.D)
3. In case the researcher decides to use APA style for documenting resources in the text, references must be placed immediately after the quote in the following order, surname of the author, year of publication, page number.
4. Opaque terms or expressions are to be explained in endnotes. List of endnotes should be placed before the list of references and resources

   Note: for more information about using APA style for documenting please check the following link:

   http://journals.qou.edu/recources/pdf/apa.pdf

Five: Peer Review & Publication Process:

All research papers are forwarded to a group of experts in the field to review and assess the submitted papers according to the known scientific standards. The paper is accepted after the researcher carries out the modifications requested. Opinions expressed in the research paper solely belong to their authors not the journal. The submitted papers are subject to initial assessment by the editorial board to decide about the eligibility of the research and whether it meets the publication guidelines. The editorial board has the right to decide if the paper is ineligible without providing the researcher with any justification.

The peer review process is implemented as follows:

1. The editorial board reviews the eligibility of the submitted research papers and their compliance with the publication guidelines to decide their eligibility to the peer review process.
2. The eligible research papers are forwarded to two specialized Referees of a similar rank or higher than the researcher. Those Referees are chosen by the editorial board in a confidential approach, they are specialized instructors who work at universities and research centers in Palestine and abroad.
3. Each referee must submit a report indicating the eligibility of the research for publication.
4. In case the results of the two referees were different, the research is forwarded to a third referee to settle the result and consequently his decision is considered definite.
5. The researcher is notified by the result of the editorial board within a period ranging from three to six months starting from the date of submission. Prior to that, the researcher has to carry out the modifications in case there are any.

Six: Scientific Research Ethics:

The researcher must:

1. Commit to high professional and academic standards during the whole process of conducting research papers, from submitting the research proposal, conducting the research, collecting data, analyzing and discussing the results, and to eventually publishing the paper. All must be conducted with integrity, neutralism and without distortion.

2. Acknowledge the efforts of all those who participated in conducting the research such as colleagues and students and list their names in the list of authors, as well as acknowledging the financial and morale support utilized in conducting the research.
3. Commit to state references soundly, to avoid plagiarism in the research.
4. Commit to avoid conducting research papers that harm humans or environment. The researcher must obtain in advance an approval from the University or the institutions he/she works at, or from a committee for scientific research ethics if there is any, when conducting any experiments on humans or the environment.
5. Obtain a written acknowledgement from the individual/individuals who are referred to in the research, and clarify to them the consequences of listing them in the research. The researcher has also to maintain confidentiality and commit to state the results of his/her research in the form of statistical data analysis to ensure the confidentiality of the participating individuals.

Seven: Intellectual Property Rights:

1. The editorial board confirms its commitment to the intellectual property rights
2. Researchers also have to commit to the intellectual property rights.
3. The research copyrights and publication are owned by the Journal once the researcher is notified about the approval of the paper. The scientific materials published or approved for publishing in the Journal should not be republished unless a written acknowledgment is obtained by the Deanship of Scientific Research.
4. Research papers should not be published or republished unless a written acknowledgement is obtained from the Deanship of Scientific Research.
5. The researcher has the right to accredit the research to himself, and to place his name on all the copies, editions and volumes published.
6. The author has the right to request the accreditation of the published papers to himself.

Contents

Machine Learning and Resampling Techniques for Enhancing Credit Card Fraud Detection in Imbalanced Dataset

Prof. Yousef Abuzir[1]*, Dr. Saleh Y. Abuzir[2]

1Computer Information Systems, Al-Quds Open University, Salfeet, Palestine.

Oricd No: 0000-0002-1220-1411
Email: yabuzir@qou.edu

2 Civil and Environmental Engineering, Brescia University, Brescia, Italy

Oricd No: 0000-0002-9825-3473
Email: s.abuzir@studenti.unibs.it

Received: 16/07/2023

Revised: 16/07/2023

Accepted: 14/10/2023

***Corresponding Author:** yabuzir@qou.edu

Citation: Abuzir, Y., & Abuzir, S. Y. Machine Learning and Resampling Techniques for Enhancing Credit Card Fraud Detection in Imbalanced Dataset. Palestinian Journal of Technology and Applied Sciences (PJTAS), 1(7). https://doi.org/10.33977/2106-000-007-001

Abstract

Objectives: This research addresses the challenge of creditcard fraud detection, complicated by highly imbalanced data where only a small fraction of transactions are fraudulent. It evaluates machine learning methods, including the Baseline Model, Logistic Regression, and Decision Tree, in conjunction with resampling techniques for handling imbalanced data in fraud detection.

Methods: The study utilizes a structured approach involving a dataset, machine learning algorithms, and resampling techniques (Oversampling, Undersampling, SMOTE) to address class imbalance in credit card fraud detection. It aims to improve accuracy by comparing models and assessing the impact of resampling methods on fraud detection performance.

Results: The results indicate that the Synthetic Minority Over-sampling Technique (SMOTE) outperforms traditional methods, achieving an accuracy of 99.89%. The Decision Tree model excels further, with 99.92% accuracy, higher recall (78.79%), and precision (98.11%). These findings underscore the potential of specialized machine learning techniques in improving fraud detection.

In conclusion, this research emphasizes the importance of resampling methods in addressing imbalanced data in credit card fraud detection. The Decision Tree model and SMOTE technique offer practical solutions for real-world applications. This study provides insights for enhancing fraud detection and highlights the role of advanced machine learning in combating credit card fraud effectively in a concise 200-word summary.

Keywords: Credit Card Fraud Detection, imbalanced data, machine learning, resampling techniques, SMOTE technique, baseline model, logistic regression, decision tree.

تقنيات التعلم الآلي وإعادة تشكيلها لتعزيز اكتشاف الاحتيال في بطاقات الائتمان في مجموعات البيانات غير المتوازنة

أ. د. يوسف أبو زر[1]*، د. صالح أبو زر[2]

1نظم الحاسوب، التكنولوجيا والعوم التطبيقية، جامعة القدس المفتوحة، سلفيت، فلسطين.

2الهندسة المدنية والبيئة، كلية الهندسة، جامعة بريشيا، بريشيا، ايطاليا.

الملخص

الملخص: الأهداف: يتناول هذا البحث التحدي المتمثل في الكشف عن الاحتيال في بطاقات الائتمان، والذي يعقده البيانات غير المتوازنة للغاية حيث يكون جزء صغير فقط من معاملات الاحتيال. البحث يناقش أساليب التعلم الآلي، بما في ذلك النموذج الأساسي والانحدار اللوجستي وشجرة القرار، جنبا إلى جنب مع تقنيات إعادة التشكيل للتعامل مع البيانات غير المتوازنة في اكتشاف الاحتيال.

الطرق: تستخدم الدراسة منهجا منظما يتضمن مجموعة بيانات، وخوارزميات التعلم الآلي، وتقنيات إعادة العينات Oversampling، Undersampling، SMOTE لمعالجة عدم التوازن الطبقي في اكتشاف الاحتيال على بطاقات الائتمان. ويهدف إلى تحسين الدقة من خلال مقارنة النماذج وتقييم تأثير طرق إعادة العينات على أداء اكتشاف الاحتيال.

النتائج: تشير النتائج إلى أن تقنية أخذ العينات الزائدة للأقلية الاصطناعية (SMOTE) تتفوق على الطرق التقليدية، حيث حققت دقة قدرها 99.89%. يتفوق نموذج شجرة القرار بشكل أكبر، مع دقة تبلغ 99.92%، واستدعاء أعلى (78.79%)، ودقة (98.11%). تؤكد هذه النتائج على إمكانات تقنيات التعلم الآلي المتخصصة في تحسين اكتشاف الاحتيال.

في الختام، يؤكد هذا البحث على أهمية أساليب إعادة المعاينة في معالجة البيانات غير المتوازنة في الكشف عن الاحتيال في بطاقات الائتمان. يقدم نموذج شجرة القرار وتقنية SMOTE حلولاً عملية لتطبيقات العالم الحقيقي. توفر هذه الدراسة رؤى لتعزيز اكتشاف الاحتيال وتسلط الضوء على دور التعلم الآلي المتقدم في مكافحة الاحتيال على بطاقات الائتمان بشكل فعال في ملخص موجز مكون من 200 كلمة.

الكلمات المفتاحية: كشف الاحتيال على بطاقات الائتمان، البيانات غير المتوازنة، التعلم الآلي، تقنيات إعادة العينات، تقنية SMOTE، النموذج الأساسي، الانحدار اللوجستي، شجرة القرار.

Introduction

The problem of credit card fraud detection is a challenge for financial institutions and customers worldwide (Kalid et al., 2020). Thus, it is essential for credit card companies to detect and prevent Credit card fraudulent transactions to protect their customers from unauthorized charges and to maintain confidence in the financial system.

The highly imbalanced nature of credit card transaction data, legitimate transactions vastly outnumber fraudulent transactions, and this leads to severe class imbalance (Botchey et al., 2020). This poses a major problem when dealing with traditional machine learning algorithms, which are typically designed to perform well on balanced datasets. Imbalanced data can impede the learning process and lead to biased models that favor the majority class, consequently hampering the accurate identification of fraudulent cases.

To address this challenge, this study focuses on applying specialized machine learning models and approaches for predicting fraud in imbalanced data. These techniques including resampling techniques, performance metric selection, Machine Learning algorithm, and synthetic sample generation (Koonsanit & Nishiuchi, 2021). In addition to resampling and performance metric evaluation, altering the algorithm can enhance fraud detection accuracy. Certain algorithms are better suited for handling imbalanced data (Khan & Malim, 2023).

This article explores the advancements in machine learning models in predicting credit card fraud detection in imbalanced data. It investigates the effectiveness of various resampling techniques, emphasizes the importance of performance metric selection, discusses Machine Learning algorithm. Real-world case studies and experimental results will be presented to illustrate the benefits and limitations of these techniques in enhancing fraud detection performance using logistic regression, and Decision tree model.

By implementing these innovative approaches, credit card companies can significantly improve their ability to detect fraudulent transactions, safeguard their customers from financial losses, and maintain confidence in credit card dealings (P & A, 2022).

Literature Survey

In recent years, studies have shown that rapid advances in machine learning have greatly improved the use of machine learning algorithms in credit card fraud detection. However, fraud detection techniques in credit card transactions remains challenging for several reasons. One significant challenge is that fraudsters continuously devise strategies to make their fraudulent transactions appear legitimate (P & A, 2022). This deceptive nature of fraudulent transactions makes it difficult for conventional machine learning algorithms to accurately identify them (Mabani et al., 2022). Additionally, credit card transaction datasets often exhibit a severe class imbalance, with a majority of legitimate transactions compared to a small number of fraudulent ones. This class imbalance poses a challenge for many Machine Learning (ML) models, because they tend to prioritize accuracy over the majority class and may ignore the minority class, which consists of fraudulent transactions (Salekshahrezaee et al., 2023).

In order to find the best algorithm for accurate detection of credit card fraud, researchers made progress in predicting imbalanced data using machine learning methods. To achieve this goal, supervised machine learning methods have been investigated. The researchers stress the importance of machine learning techniques to address the complex problems associated with data-based fraud detection in credit cards. For instance, the FraudMiner model was proposed to address highly imbalanced and anonymized credit card transaction instances. The model studies the problem of the class imbalance by employing frequent itemset mining to uncover patterns for both legal and fraudulent transactions on an individual customer level (Mohammed, 2022). By considering individual customer behavior, the model aims to enhance fraud detection accuracy beyond relying solely on aggregated transaction data. Another study focused on selecting the best credit card fraud detection techniques for utilization (Manek et al., 2019).

The study of Salekshahrezaee (Salekshahrezaee et al. 2023), considered a highly imbalanced dataset, and applied the different types of machine learning techniques (supervised and unsupervised) to detect credit card fraudulent transactions in credit cards. The results presented in their conclusion that unsupervised machine learning algorithms outperformed supervised ones in handling the skewness caused by the class imbalance.

These advancements in imbalanced data prediction using machine learning techniques are crucial for improving fraud detection in credit card transactions (Tran et al., 2019). By leveraging different machine learning algorithms and techniques, researchers are striving to achieve higher accuracy in identifying fraudulent transactions. Credit card fraud detection poses significant challenges due to fraudsters' evolving techniques to make their transactions appear legitimate and the inherent class imbalance in credit card dataset (Kumar, 2021). This suggests addressing issues with highly balanced datasets is a major concern (Hamid et al., 2022). As legitimate credit card transactions vastly outnumber fraudulent ones, many machine learning algorithms struggle to achieve satisfactory performance (Mabani et al., 2022), (Hasanin et al., 2019).

To tackle these challenges, (Hasanin et al., 2019) have focused on developing algorithms capable of effectively handling imbalanced data. One such model is FraudMiner, which employs frequent itemset mining to identify legal and fraudulent transaction patterns for each customer in highly imbalanced and anonymous credit card data transaction (Seeja & Zareapoor, 2014), Another challenge in credit card fraud detection using machine learning is fraudsters adopt advanced and highly accurate techniques to make fraudulent transactions in credit cards appear legitimate (Fang et al., 2021).

Several studies have explored the use of automated learning algorithms for credit card fraud detection. For example, Mairaj et al., (Mairraj et al. 2019) studied the applicability of machine learning techniques such as decision trees (DT), logistic regression (LR), and random forests (RF), to identify legal and fraudulent transaction. The study also discussed the challenges associated with imbalanced data and the use of resampling techniques to solve these challenges. Additionally, Meera E (Meera E, 2022) evaluated the use of decision trees, random forests, and logistic regression for detecting credit card fraud. In order to enhance the performance of machine learning models, the study stressed the significance of feature selection and hyperparameter adjustment. The study stressed the need of feature selection and hyperparameter tuning in order to improve the performance of machine learning models. Kwaku et al. (2023) evaluated how well three different machine learning algorithms logistic regression, decision trees, and random forest performed at detecting fraudulent credit card transactions. Also, the same work assessed by Kwaku et al. (Kwaku et al. 2023). In order to improve the performance of the machine learning models, the study underlined the significance of attribute selection and hyperparameter approaches.

Data And Methodology

In this section, we discuss the dataset, the different machine learning algorithms and the resampling techniques such as Oversampling from the minority, undersampling from the majority and SMOTE. The methodology of this research using machine learning techniques follows a structured approach to detect fraudulent transactions. The diagram in Figure 1 shows an overview of the most important stages and the various sub-steps in the study.

Overall, this methodology provides a structured and comprehensive approach to tackle the challenges posed by class imbalance and improve the accuracy of fraud detection in credit card transactions. By comparing different models and utilizing resampling techniques, this study aims to identify the most suitable model and evaluate the impact of these techniques on enhancing fraud detection performance.

Modeling Approach

The system architecture is designed to comprehensively evaluate the performance of different machine learning models and resampling techniques in the context of credit card fraud detection. It consists of the following sequential steps:

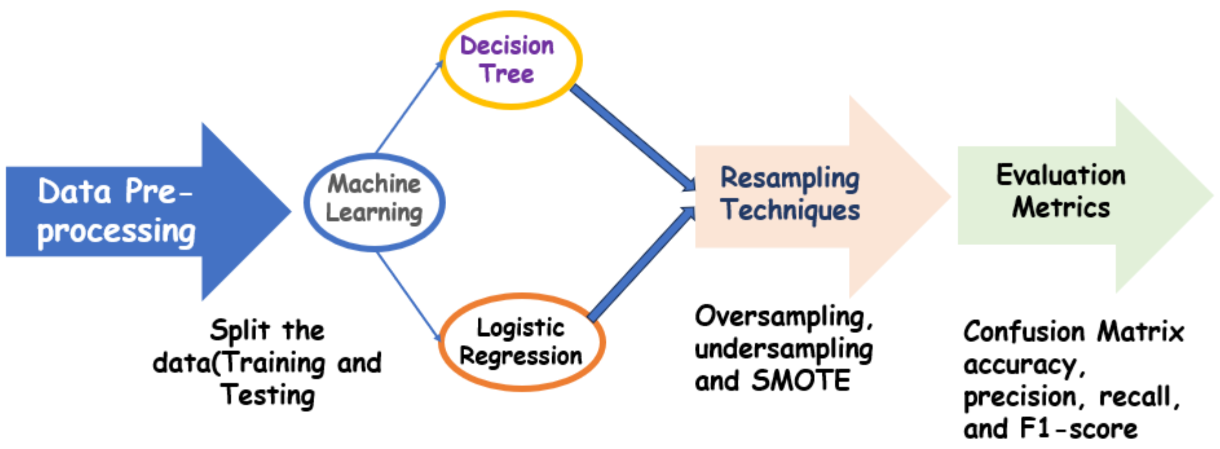

Figure 1: The System Architecture - Block Diagram

- Data Preprocessing: The initial step involves data cleaning, transformation, and handling missing values to prepare the dataset for analysis. This ensures that the data is in a suitable format for further processing.
- Data Exploration: plays a pivotal role. It involves understanding data distribution, analyzing features, and employing visualizations to uncover patterns and anomalies. This step ensures that subsequent decisions regarding model selection, resampling techniques, and feature engineering are well-informed and tailored to effectively address the challenges posed by class imbalance.
- Data Splitting: The preprocessed data is divided into two subsets: a training set and a testing set. The training set is used to train the machine learning models, while the testing set is reserved for evaluating their performance.
- Machine Learning Models: Two machine learning models, Decision Tree and Logistic Regression, are applied to the training data separately. Each model learns from the training data to make predictions about fraudulent and non-fraudulent transactions.
- Resampling Techniques: To address the class imbalance problem inherent in fraud detection datasets, three resampling techniques are employed: Oversampling, Undersampling, and Synthetic Minority Over-sampling Technique (SMOTE). Each of these techniques aims to create a more balanced dataset for model training and evaluation.
- Model Evaluation Metrics: The performance of each machine learning model, as well as the resampling techniques, is assessed using a comprehensive set of evaluation metrics. These metrics include the Confusion Matrix, Accuracy, Precision, Recall (Sensitivity), and F-1 Score. The Confusion Matrix provides a detailed breakdown of correct and incorrect predictions, while the other metrics offer insights into various aspects of model performance.

This systematic approach ensures a thorough evaluation of different techniques and models, allowing for a comprehensive comparison of their effectiveness in detecting credit card fraud. The results and discussion sections are subsequently separated to provide a clear and structured presentation of findings without integration, reducing reader confusion.

Dataset

The dataset of 284,807 transactions prepared by European cardholders in September 2013 were used in this research. The fraud transactions represent a very small percentage of 0.172% of all the transactions in the dataset (only 492 transactions).

For confidentiality reasons, European cardholders applied Principal Component Analysis (PCA), which masks the original features of the data. However, some important features like time and amount are accessible. Table 1 presents the available features:

Table 1 Description of the features

Feature	Description
V1-V28	Principal components resulting from PCA transformation
Time	Time elapsed (in seconds) since the first transaction in the dataset
Amount	Transaction amount
Class	Response variable: 1 for fraud transactions, 0 otherwise

In preparing the dataset, different functions were used like, isnull() to find the null values in the dataset, the sum() function used to calculate the total number of null values in each column, and as confirmed by the df.info() and df.head() functions, the dataset contains 284,807 entries and 31 columns.

By analyzing the pie chart in figure 2, we observe that the fraudulent data accounts for only 0.17% of the entire dataset.

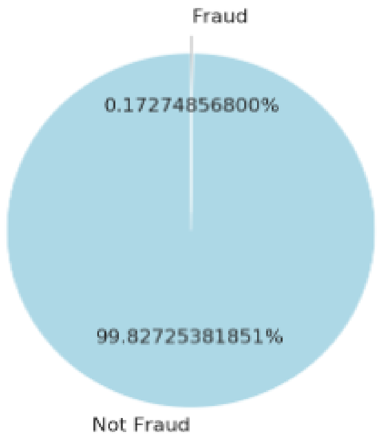

Figure 2 pie chart for fraudulent data- distribution of transactions

This observation is logical because:

It indicates that the credit card company has been successful in preventing fraud, as the proportion of fraudulent transactions is extremely low 0.17%.

Fraudulent transactions are inherently rare occurrences, which aligns with the low percentage mentioned in point (1).

However, this significant class imbalance poses challenges for predictive modeling. The imbalanced nature of the data can impact our data preprocessing techniques and the selection of appropriate modeling algorithms, evaluation metrics, and resampling strategies.

The observation underscores the success of the credit card company in preventing fraud and the infrequent nature of these occurrences. However, it's vital to clarify that our study does not diminish the importance of detecting rare fraud cases. Instead, our research addresses the challenges presented by class imbalance in predictive modeling. Recognizing these rare events is critical, as even a small percentage of fraudulent transactions can lead to significant financial losses and risks. Our study emphasizes the necessity for specialized techniques and models tailored to handle such scenarios effectively. Our focus is on developing and evaluating these techniques to enhance fraud detection, ensuring accurate identification even in rare cases.

The dataset's class imbalance emphasizes the necessity for specialized methods and assessment metrics to handle the problems that class unbalanced data in credit card fraud detection presents.

The generated heatmap in Figure 3 shows the correlation values between all the variables in the dataset, excluding the 'Class' variable. Heatmaps can visually represent correlation strength and direction using color shading, with cooler colors indicating negative correlation and warmer colors indicating positive correlation. Numeric values within each cell display the precise correlation coefficient for the corresponding variables.

Results And Discussion

This section presents a comprehensive analysis of the results, performance metrics of different models such as accuracy, recall, precision, specificity, and F1 score. This section will further discuss the implications of the findings. A comparison of several approaches, including resampling techniques and machine learning models like logistic regression and decision trees, will also be covered to show their advantages and disadvantages. The knowledge collected from this analysis will advance knowledge of fraud detection techniques and offer helpful direction for next studies and the creation of effective fraud detection systems.

1-logistic regression

A statistical model known as logistic regression is utilized for binary classification tasks and is characterized by its simplicity and interpretability. Based on the supplied features, it calculates the likelihood that a particular instance will belong to a particular class.

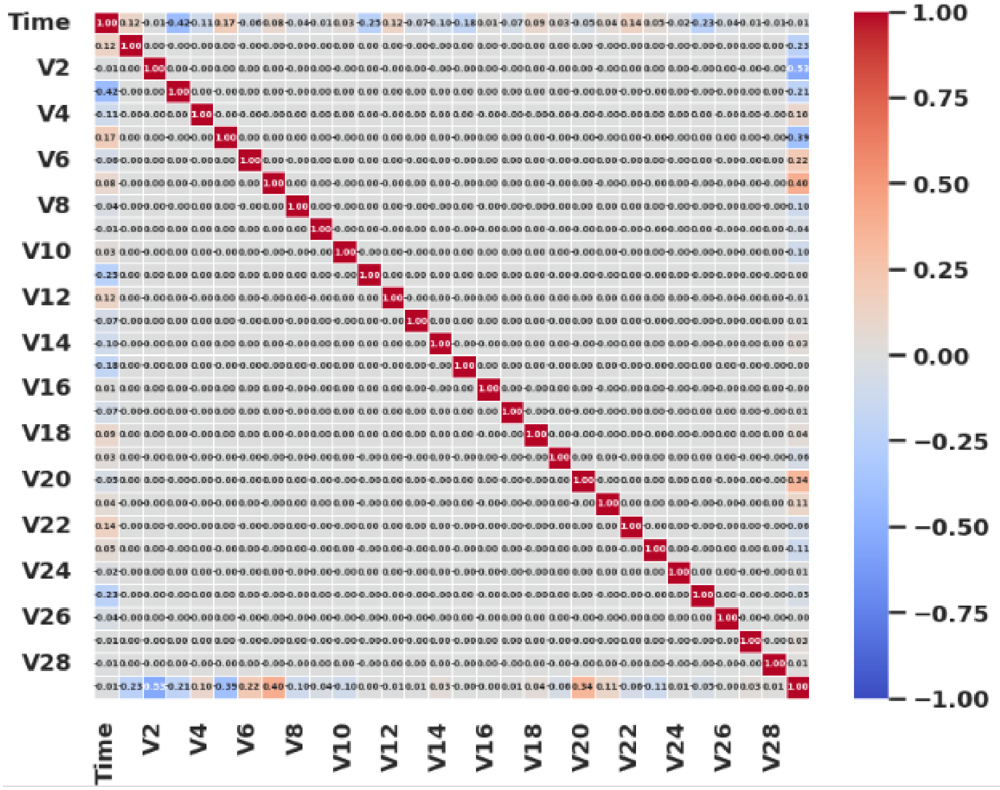

Figure 3 The heatmap shows the correlation values between all the variables in the dataset

Logistic regression is a theoretical model that uses the sigmoid function to assign real-valued numbers to probabilities between 0 and 1. By applying this function to a linear set of input features using coefficients or weights, logistic regression predicts the probability of an event (Abuzir Y., 2018). The model learns these coefficients from the training data through an optimization process called maximum likelihood estimation.

The logistic regression equation can be represented as follows:

$$p = 1 / (1 + e^{(-z)})$$

where p is the probability of the event, z is the linear combination of the input features and their corresponding coefficients. The coefficients are determined during the model training process.

It is crucial to pick suitable performance metrics that offer deeper insights when assessing our model's performance on imbalanced data sets. A table displaying both the kinds of accurate predictions the model made and its confusion matrix are shown in figure 4. The following indicators are suggested for credit card fraud detection, where the goal is to accurately categorize fraud situations (Table 2).

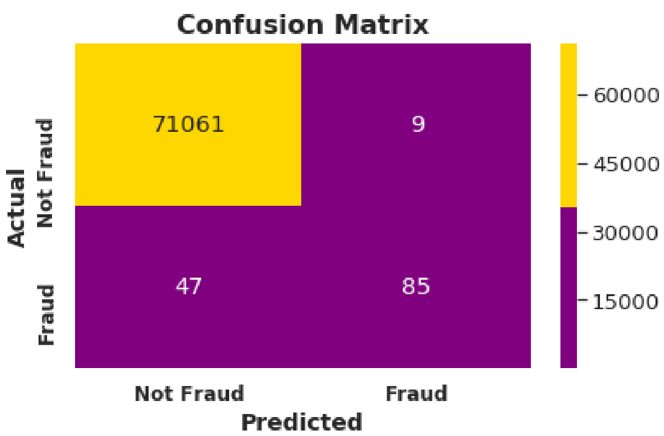

Figure 4 Confusion matrix for logistic regression performs on this dataset

Table 2 Performance Metrics for Baseline Model and Logistic Regression

Performance Metrics for	Accuracy	Recall	Precision	Specificity	F1 Score
Logistic Regression	99.89%	64.47%	90.43%	99.99%	75.22%

The analysis highlighted the significance of prioritizing the recall score and taking into account specific objectives in credit card fraud detection. Maximizing the identification of fraudulent transactions is essential, and further ways could be investigated, such as resampling techniques (under- or over-sampling), or more sophisticated algorithms created for imbalanced datasets.

2-Decision trees

To address the limitations of our current logistic regression model on the imbalanced dataset, we can explore alternative algorithms that may perform better. One such algorithm is the decision tree, which has shown promise in handling imbalanced data.

Decision trees operate by constructing a hierarchical structure similar to if/else structure to make predictions. Decision trees model can be a good selection for imbalanced datasets as it allows both classes to be considered and addressed during the learning process (Abuzir Y., Abuzir S. Y., 2020).

By implementing a decision tree algorithm, we can potentially improve the model's ability to detect fraudulent transactions. The decision tree can capture nonlinear relationships and interactions between features, which may be important for identifying fraud patterns that are not captured by logistic regression.

By trying different algorithms and exploring decision trees in particular, we can leverage their inherent capabilities to better handle the class imbalance in our dataset and improve the detection of fraudulent transactions.

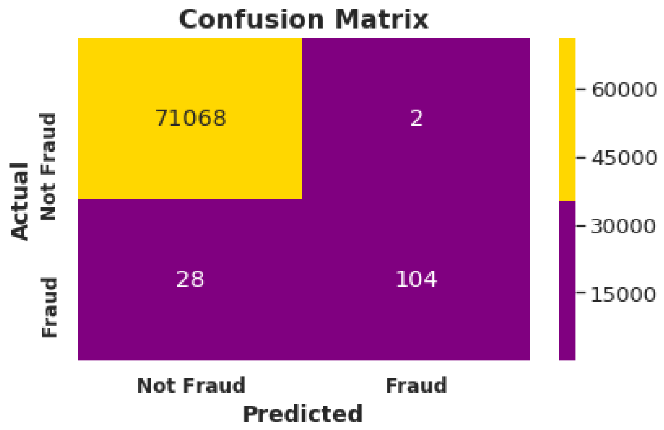

Figure 5 Confusion matrix for Decision Tree

Table 3 Performance Metrics for Baseline Model, Logistic Regression and Decision Tree

Performance Metrics for	Accuracy	Recall	Precision	Specificity	F1 Score
Logistic Regression	99.89%	64.47%	90.43%	99.99%	75.22%
Decision Tree	99.92%	78.79%	98.11%	99.99%	87.58%

The confusion matrix (figure 5) and performance metrics (Table 3) indicate that Decision Tree model correctly predicted a large number of not fraud cases (actual not fraud) with a high specificity of 99.99%. The precision of 98.11% indicates that when the model predicts a transaction as fraud, it is correct 98.11% of the time. The overall accuracy of the model is 99.92%, indicating that Decision Tree model has a highest prediction of other models.

Compared to the decision tree model, which had an accuracy score of 0.98, the logistic regression model had a superior score of 99.9%. This shows that both models illustrate the overall accurate predictions. The most trustworthy criterion to assess model performance on unbalanced datasets may not, however, be accuracy alone.

When we look at the performance metrics F1 score, which depends on both accuracy and recall, we find that the decision tree model performs better than the logistic regression model. The decision tree model has a value of 0.87 for F1 score,

indicating a better balance between precision and recall, while the logistic regression model has a value less than 0.752 to F1 score.
Furthermore, the decision tree model has a higher recall score of 0.79 compared to the logistic regression model with a recall score of 0.64, which assesses the model's capacity to accurately detect positive cases (fraudulent transactions). The decision tree methodology is therefore more effective at catching and identifying fraudulent transactions, according to this data.
While both models exhibit excellent accuracy, it is evident that the decision tree model performs better in terms of F1 score and recall score. As a result, it can be inferred that the decision tree model is more successful in managing the class imbalance and detecting fraudulent transactions. Figure 6 visualizes the relationship between accuracy, F1 score, and recall score.

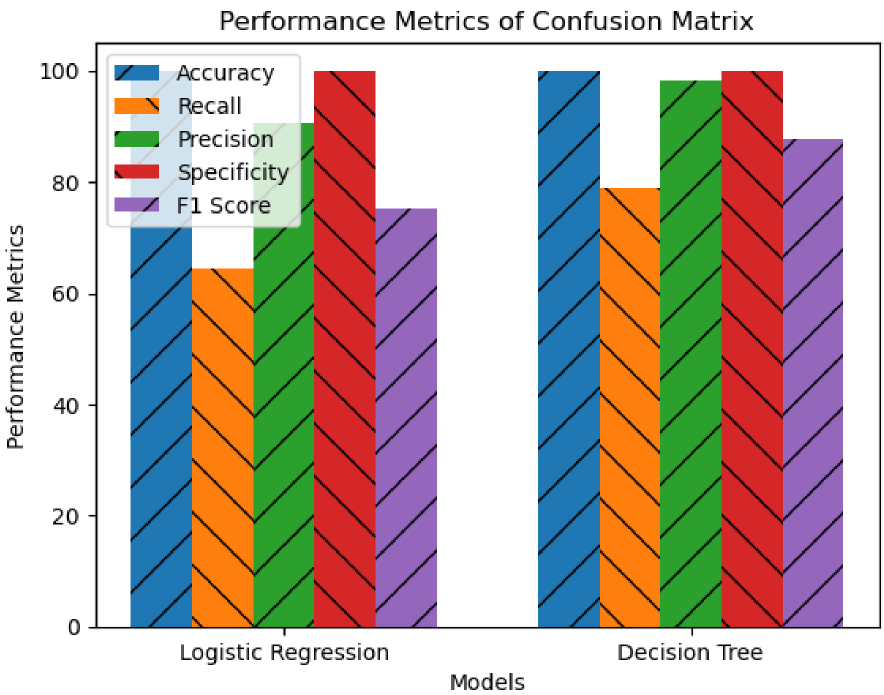

Figure 6 relationship between accuracy, F1 score, and recall score

The comparison of the two models reveals that the performance of Decision Tree model is the best in terms of recall, precision, and F1 score. It effectively detects a greater percentage of real fraud situations while exhibiting a low false positive rate. A good balance between precision and recall is achieved by the Logistic Regression model.
3-Resampling Techniques
Dealing with imbalanced datasets is a common challenge in machine learning, where the distribution of classes is highly skewed, with one class being significantly more general than the other(s). Due to their propensity to favor the majority class and ignore the minority class, machine learning models may perform poorly as a result of this class imbalance (Wu et al., 2021). The problem of class imbalance in datasets has given rise to a number of techniques.
Several techniques have been developed to address the issue of class imbalance in dataset. These technologies work to balance the data collection by increasing the representation of the minority category (oversampling) or reducing the representation of the majority category (undersampling) or using hybrid techniques that combine oversampling and undersampling methods to achieve a more balanced data set. One of the common methods is the technique of excessive artificial minority (SMOTE) (Liu et al., 2022).

1. Oversampling the Minority Class: Oversampling is the process of raising the minority class's number of instances to equal the majority class's number of instances. This can be accomplished by creating synthetic samples based on current instances or by randomly reproducing existing examples from the minority class (Rajendran et al., 2020). The objective is to give the model more cases of the minority class so that it may learn the patterns and get better at correctly classifying instances of the minority class (Wu et al., 2021).
2. Undersampling the Majority Class: Undersampling attempts to balance a dataset by reducing the number of instances in the majority class. This can be done by arbitrarily eliminating instances from the class that makes up

the majority until the required balance is attained. Undersampling enables the model to pay more attention to the minority class by decreasing the dominance of the majority class, which helps avoid the model from being biased towards forecasting the majority class (Akbar et al., 2020) (Ba-Rukab et al., 2022).

3. SMOTE (Synthetic Minority Over-sampling Technique): SMOTE is a is a hybrid resampling approach that combines oversampling and artificial sample generation. SMOTE develops synthetic samples by interpolating between existing minority class instances rather than merely replicating examples from the minority class. It locates each minority class instance's closest neighbors and creates new instances along the line segments jogging between them. The risk of overfitting is reduced by this method's introduction of diversity to the synthetic samples (Akondi et al., 2022).

These resampling techniques provide different ways to tackle class imbalance and improve the performance of machine learning models on imbalanced datasets. However, the selection of resampling technology depends on the description of the problem, the data group and algorithm used. Researchers are advised to try different techniques and evaluate their performance on the model using appropriate evaluation metrics.

The Scikit-Learn resampling module is used in this section to study various resampling methods. When one class is underrepresented disproportionately in machine learning tasks compared to the other(s), resampling approaches are frequently used to solve the class imbalance. These methods include using the Synthetic Minority Over-sampling Technique (SMOTE), undersampling the majority class, and oversampling the minority class.

We'll utilize Scikit-Learn's resampling module, which enables us to randomly reproduce samples from the minority class, to carry out oversampling.

Let's examine the confusion matrix (Figure 7) for the oversampling dataset of Fraud Detection in Credit Cards to interpret and explain it.

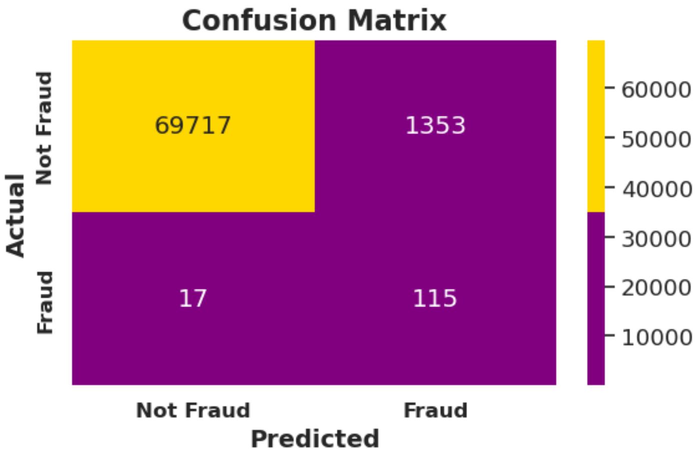

Figure 7 Confusion matrix for Oversampling

Table 4 Metrics oversampling technique applied to the minority class in credit card fraud detection

Performance Metrics for	Accuracy	Recall	Precision	Specificity	F1 Score
Oversampling Minority Class	98.17%	87.12%	7.86%	98.09%	14.34%

The oversampling technique applied to the minority class in credit card fraud detection resulted in the following metrics (Table 4). The given confusion matrix (Figure 7) and associated metrics (Table 4) show that the oversampling technique improved the model's ability to identify fraud cases, resulting in a high recall and specificity. However, the low precision indicates a higher rate of false positives, suggesting room for improvement in correctly classifying non-fraud cases. The F1 score provides a comprehensive evaluation, considering both precision and recall in a single metric.

The confusion matrix (Figure 8) for Undersampling summarizes the model's predictions in relation to the data's actual labels. We may evaluate the model's performance and obtain insight into the kinds of errors it is making by evaluating the numbers in the matrix.:

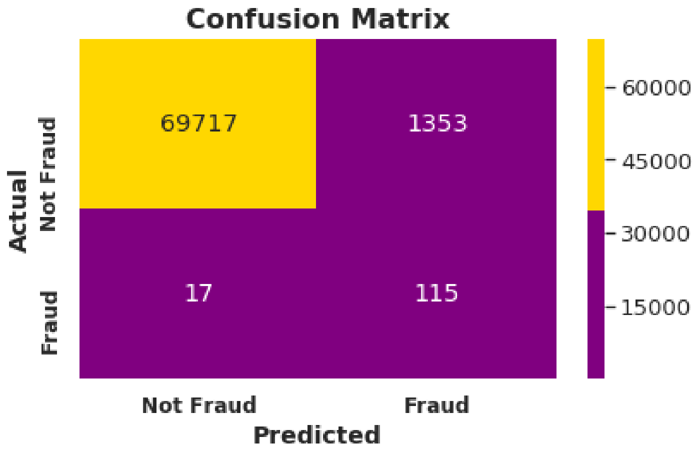

Figure 8 Confusion Matrix Undersampling

Table 5 Metrics Undersampling technique applied to the minority class in credit card fraud detection

	Accuracy	Recall Score	Precision Score	Specificity	F1 Score
Undersampling Majority Class	97.45%	86.36%	6.26%	97.61%	11.53%

Table 5 shows the performance metrics of the undersampling technique applied to the minority class in credit card fraud detection.
Figure 9 presents the confusion matrix for the SMOTE (Synthetic Minority Over-sampling Technique), let's analyze the table in figure 8.

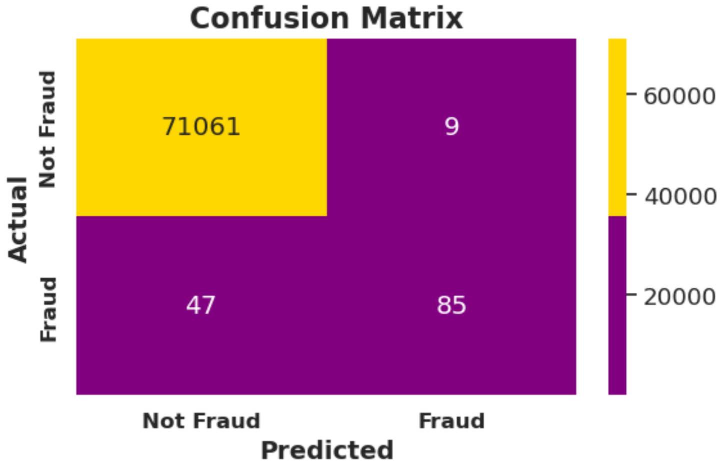

Figure 9 Confusion Matrix SMOTE (Synthetic Minority Over-sampling Technique)

In summary, the confusion matrix and associated metrics (Table 6) for the SMOTE dataset show that the model has a high overall accuracy, accurately classifying both fraud and non-fraud cases. The model exhibits a high precision, indicating a low rate of false positives. However, the recall is moderate, suggesting that the model identifies a moderate proportion of the fraud cases. The specificity is high, indicating a very low rate of false positives in the non-fraud cases. The F1 score reflects a balance between precision and recall, providing a single metric to evaluate the model's performance.

Table 6 Metrics SMOTE technique applied to the minority class in credit card fraud detection

	Accuracy	Recall Score	Precision Score	Specificity	F1 Score
SMOTE	99.89%	64.47%	90.43%	99.99%	75.22%

The SMOTE technique demonstrated excellent performance in credit card fraud detection. It achieved a high accuracy 99.89% and specificity 99.99%, indicating a low rate of false positives. The precision score 90.43% was also high, indicating a low rate of false positives. However, the recall score 64.47% was relatively moderate, suggesting that there is room for improvement in identifying more fraud cases. Figure 10 shows a comparison between the different sampling techniques.

In summary, the SMOTE technique outperformed the undersampling and oversampling techniques in terms of accuracy, precision, specificity, and F1 score. However, it had a lower recall score compared to oversampling. The SMOTE technique effectively addressed the class imbalance issue, providing a more balanced representation of both classes and achieving better overall performance in detecting fraudulent transactions.

Table 7 summarizes the performance metrics for all techniques.

	Accuracy	Recall Score	Precision Score	Specificity	F1 Score
Decision Tree without Resampling	99.92%	78.79%	98.11%	99.99%	87.58%
Logistic Regression without Resampling	99.89%	64.47%	90.43%	99.99%	75.22%
Oversampling Minority Class	98.17%	87.12%	7.86%	98.09%	14.34%
Undersampling Majority Class	97.45%	86.36%	6.26%	97.61%	11.53%
SMOTE	99.89%	64.47%	90.43%	99.99%	75.22%

As shown in table 7, the decision tree model without resampling demonstrated the highest performance across all metrics, achieving high accuracy, recall, precision, specificity, and F1 score. The logistic regression model without resampling and the SMOTE technique also showed competitive performance. However, the oversampling and undersampling techniques exhibited lower precision and F1 scores, indicating a higher number of false positive predictions. The choice of technique should be based on the desired balance between different evaluation metrics for fraud detection in credit card transactions.

Due to the following characteristics, the decision tree model without resampling was able to leverage the inherent strengths of the algorithm and achieve high accuracy, recall, precision, specificity, and F1 score: Handling Imbalanced Data, Feature Importance, Non-Linear Relationships, and Flexibility in Decision Rules. It's crucial to remember that the decision tree model's performance can change depending on the dataset and the precise details of the current fraud detection situation.

Conclusion

Credit card fraud detection is a challenging task due to the highly imbalanced nature of the data, where fraud cases account for a relatively small percentage of all transactions. Traditional machine learning algorithms tend to favor the majority class and fail to handle imbalanced data well.

The best algorithm founded in this study is Decision Tree without resampling achieved an accuracy of 99.92%, indicating a high overall correct classification rate. It demonstrated a recall score of 78.79%, meaning it successfully identified approximately 78.79% of the actual fraud cases. The precision score of 98.11% suggests that when the model predicted a transaction as fraud, it was correct around 98.11% of the time. The specificity value of 99.99% indicates a low false positive rate, accurately identifying non-fraudulent transactions. The F1 score of 87.58% reflects a good balance between the two scores precision and recall.

Logistic Regression (without Resampling) achieved a slightly lower accuracy of 99.89% compared to the decision tree. It showed a recall score of 64.47%, indicating that it identified approximately 64.47% of the actual fraud cases. The precision score of 90.43% suggests that when the model predicted a transaction as fraud, it was correct around 90.43% of the time. The specificity value of 99.99% indicates a low false positive rate. The F1 score of 75.22% indicates a reasonable balance between precision and recall.

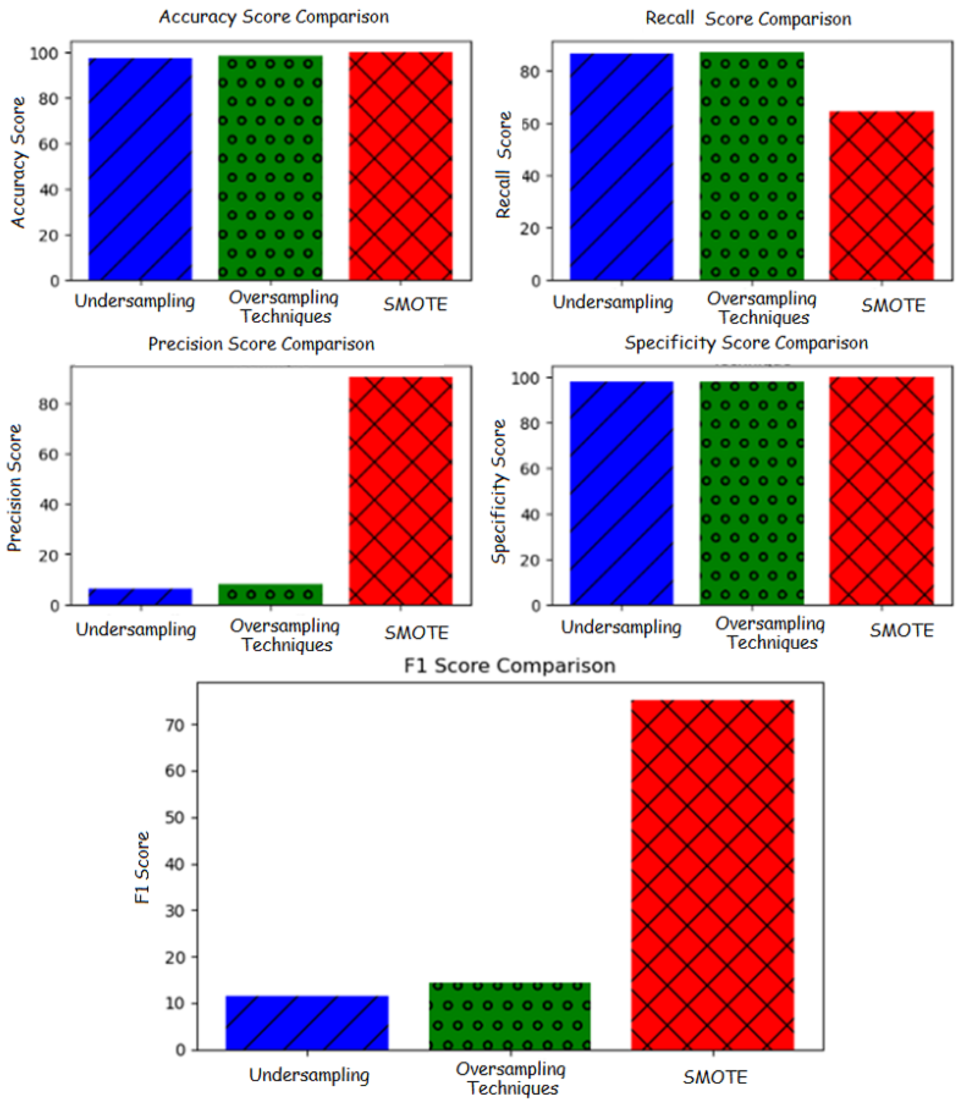

Figure 10 A Comparison between the different sampling techniques

The oversampling technique achieved an accuracy of 98.17%. It showed a higher recall score of 87.12%, indicating that it successfully identified a larger proportion (87.12%) of the actual fraud cases compared to the logistic regression model. However, because of the low precision score of 7.86%, there were a lot of false positive predictions. The specificity value of 98.09% suggests a moderate false positive rate. The F1 score of 14.34% indicates a trade-off between precision and recall.

The undersampling technique achieved an accuracy of 97.45%. It demonstrated a recall score of 86.36%, indicating it identified a relatively high proportion (86.36%) of the actual fraud cases. However, the precision score was only 6.26%, suggesting a significant number of false positive predictions. The specificity value of 97.61% implies a moderate false positive rate. The F1 score of 11.53% indicates a trade-off between precision and recall.

SMOTE method's accuracy of 99.89% was comparable to that of both the logistic regression and decision tree models. A recall score of 64.47% was shown, matching the logistic regression model's results. With a precision score of 90.43%, fraud cases can be predicted with a respectable degree of accuracy. A low percentage of false positives is indicated by the specificity value of 99.99%. Similar to the logistic regression model, the F1 score of 75.22% indicates a balance between precision and recall.

These percentage values provide a clearer understanding of the performance of each technique in terms of their accuracy, recall, precision, specificity, and F1 score. The study highlights the importance of using appropriate techniques to handle imbalanced data in fraud detection.

Using such models in real-world applications can significantly enhance fraud prevention efforts and reduce financial losses for credit card users and bank. Machine Learning algorithms like logistic regression and decision trees can help banks identify and prevent unauthorized activities.

Further research, progress and evaluation in this field can lead to continued advancements in fraud detection algorithms, contributing to a safer and more secure financial environment. Using decision trees also opens up opportunities for techniques such as ensemble learning, where multiple decision trees are combined to make predictions.

References

- Abuzir Y. (2018). Predict the Main Factors that Affect the Vegetable Production in Palestine Using WEKA Data Mining Tool, **Palestinian Journal of Technology and Applied Sciences (PJTAS)**, 1: (58-71).
- Abuzir Y., Abuzir S. Y., (2020). Data Mining Techniques for Prediction of Concrete Compressive Strength (CCS), **Palestinian Journal of Technology and Applied Sciences (PJTAS)**, 3 (57-72).
- Akbar, A. T., Husaini, R., Akbar, B. M., & Saifullah, S. (2020). A proposed method for handling an imbalance data in classification of blood type based on Myers-Briggs type indicator. Available: https://scite.ai/reports/10.14710/jtsiskom.2020.13625.
- Akondi, V. S., Menon, V., Baudry, J., & Whittle, J. (2022). Novel Big Data-Driven Machine Learning Models for Drug Discovery Application. **Molecules**, 27(3), 594., Available: https://doi.org/10.3390/molecules27030594.
- Botchey, F. E., Qin, Z., & Hughes-Lartey, K. (2020). Mobile Money Fraud Prediction - A Cross-Case Analysis on the Efficiency of Support Vector Machines, Gradient Boosted Decision Trees, and Naïve Bayes Algorithms. ***Information***. 11(8):383. Available: https://scite.ai/reports/10.3390/info11080383.
- Fang W., Li X., Zhou P., Yan J., Jiang D. and Zhou T., (2021). Deep Learning Anti-Fraud Model for Internet Loan: Where We Are Going, in **IEEE Access**, vol. 9: (9777-9784), doi: 10.1109/ACCESS.2021.3051079.
- Hamid, M. H. A., Yusoff, M., & Mohamed, A. (2022). Survey on Highly Imbalanced Multi-class Data. **International Journal of Advanced Computer Science and Applications**, 13(6):211-229. https://doi.org/10.14569/ijacsa.2022.0130627.
- Hasanin, T., Khoshgoftaar, T. M., Leevy, J. L., & Bauder, R. A. (2019). Severely imbalanced Big Data challenges: investigating data sampling approaches. **Journal of Big Data**, 6(1), (2019, November 30). https://doi.org/10.1186/s40537-019-0274-4.
- Kalid, S. N., Ng, K. H., Tong, G., & Khor, K. (2020). A Multiple Classifiers System for Anomaly Detection in Credit Card Data with Unbalanced and Overlapped Classes, **IEEE Access**, vol. 8: 28210-28221doi: 10.1109/ACCESS.2020.2972009.
- Khan, A. K. A., & Malim, N. H. A. H. (2023). Comparative Studies on Resampling Techniques in Machine Learning and Deep Learning Models for Drug-Target Interaction Prediction. **Molecules** 28(4):1663. https://doi.org/10.3390/molecules28041663.
- Koonsanit, K., & Nishiuchi, N. (2021). Predicting Final User Satisfaction Using Momentary UX Data and Machine Learning Techniques, **J. Theor. Appl. Electron. Commer. Res.** *16*(7):3136-3156. https://doi.org/10.3390/jtaer16070171.
- Kumar, T. (2021). Comparison of Logistic Regression and Decision Tree method for Credit Card Fraud Detection, **International Journal for Research in Applied Science & Engineering Technology (IJRASET)**, 9(V):680-683. https://scite.ai/reports/10.22214/ijraset.2021.34241.
- Kwaku J. A., Tawiah K., Pels W., Sandra A. H., Dwamena H.A., Owiredu E. O., Samuel A. A., Eshun J. (2023), A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions, **Decision Analytics Journal**, 6:100-163, https://doi.org/10.1016/j.dajour.2023.100163.
- Liu, A., Cheng, L., & Yu, C. (2022). SASMOTE: A Self-Attention Oversampling Method for Imbalanced CSI Fingerprints in Indoor Positioning Systems, ***Sensors*** *22*(15), 5677; https://doi.org/10.3390/s22155677.
- Mabani, C., Tuskov, A., & Shchanina, E. (2022). Detection Of Credit Card Frauds With Machine Learning Solutions: An Experimental Approach. **Krasnoyarsk Science**, 11(3), 17-28. https://doi.org/10.12731/2070-7568-2022-11-3-17-28.
- Manek, H., Kataria, N., Jain, S., & Bhole, C. (2019, November 15). Various Methods for Fraud Transaction Detection in Credit Cards. **Journal of Ubiquitous Systems & Pervasive Networks**, 12(1): 25-30. https://scite.ai/reports/10.5383/juspn.12.01.004.

- Maniraj, S.P., Saini, A., Ahmed, S., & Sarkar, S.D. (2019) Credit Card Fraud Detection using Machine Learning and Data Science. **International Journal of Engineering Research & Technology (IJERT)**, 8(9):110-115. DOI:10.17577/ijertv8is090031
- Meera E., (2022). **Credit Card Fraud Detection Using Machine Learning**. Thesis. Rochester Institute of Technology, NY, USA.
- Mohammed, S. J. (2022). Detection and Prevention WEB-Service for Fraudulent E-Transaction using APRIORI and SVM. **Al-Mustansiriyah Journal of Science**, 33(4):72–79. https://scite.ai/reports/10.23851/mjs.v33i4.1242.
- Raghavendra R. P., Sivanesh K. A. (2022). A Fraudulent Transaction Prediction in Credit Card by Using Novel LGBA over LR Algorithms. **Advances in Parallel Computing: Algorithms, Tools and Paradigms** 41: (515-521). https://doi.org/10.3233/apc220073.
- Rajendran, K., Jayabalan, M., & Thiruchelvam, V. (2020). Predicting Breast Cancer via Supervised Machine Learning Methods on Class Imbalanced Data. **International Journal of Advanced Computer Science and Applications**, 11(8): 54-63, https://doi.org/10.14569/ijacsa.2020.0110808.
- Seeja, K. R. and Zareapoor, M. (2014). Fraudminer: a novel credit card fraud detection model based on frequent itemset mining. **The Scientific World Journal**, 1-10. https://doi.org/10.1155/2014/252797.
- Tran, L. V., Tran, T., Tran, L. T., & Mai, A. (2019). Solve fraud detection problem by using graph-based learning methods. **Journal of Engineering and Science Research**, 3(4), 28-31. https://doi.org/10.26666/rmp.jesr.2019.4.6
- Wu, C., Wang, N., & Wang, Y. (2021). Increasing minority recall support vector machine model for imbalanced data classification. **Discrete Dynamics in Nature and Society**, 1-12. https://doi.org/10.1155/2021/6647557.

https://journals.qou.edu/index.php/PJTAS
E-ISSN: 2521-411X/ P-ISSN: 2520-7431

Using Linear Mathematical Programming to Maximization the Net Return per Unit Area for Irrigated Crops (case study)

Dr. Maya Al-abdala[1], Dr. Safwan Aboassaf[2], Dr. Afraa Sallowm[3]*

1Researcher, General Commission for Scientific Agricultural Research (GCSAR), Syria.
Orcid No: 0000-0003-0370-6191
Email: mayaabdala6@gmail.com

2Researcher, General Commission for Scientific Agricultural Research (GCSAR), Syria.
Orcid No: 0009-0006-8562-178X0
Email: aboassaf2011@hotmail.com

3Assistant Professor, Faculty of Agriculture, Damascus University, Syria.
Orcid No: 0009-0003-5751-4192
Email: asallowm@yahoo.com

Received:
16/07/2023

Revised:
4/02/2023

Accepted:
14/10/2023

***Corresponding Author:**
yabuzir@qou.edu

Citation: Al-abdala, M., Aboassaf , S., & Sallowm, A. Using Linear Mathematical Programming to Maximization the Net Return per Unit Area for Irrigated Crops (case study). Palestinian Journal of Technology and Applied Sciences (PJTAS), 1(7). https://doi.org/10.33977/2106-000-007-002

Abstract

Objectives: The article aims to propose an optimal cropping pattern that maximizes net returns per unit area in Swaida Governorate, Syria.

Methods: Using mathematical linear programming, is an important tool for studying different aspects of crop structures with all the constraints they face that hinder production, such as changing weather conditions, water problems, labor problems and economic conditions. Through a questionnaire targeting the farmers of irrigated summer and winter vegetables during the 2020 season, for a sample of 106 farmers. By redistributing the area allocated to each crop

Results: , the study found an 80.8% increase in total net income compared to the actual crop pattern while reducing water consumption, amounting to approximately 5.6 million m3 of water, while the total consumption of the actual cropping by sample farmers was 5.9 million m3 of water.

Conclusions: The contributions of the paper include suggesting the optimized cropping pattern, increasing net income, reducing water consumption, and emphasizing the importance of using linear programming in agriculture decision-making. While preserving the area allocated for cultivating basic crops such as wheat, potatoes, tomatoes and watermelon.

Keywords: Crop pattern, linear mathematical programming, maximization, optimization, Swaida governorate.

استخدام البرمجة الخطية الرياضة لتعظيم عائد وحدة المساحة للتركيبة المحصولية المروية (دراسة حالة)

د. مايا العبد الله[1]، د. صفوان أبو عساف[2]، د. عفراء سلوم[3]*

[1]باحث، الهيئة العامة للبحوث العلمية الزراعية، سوريا.

[2]باحث، الهيئة العامة للبحوث العلمية الزراعية، سوريا.

[3]استاذ مساعد، قسم الاقتصاد، كلية الهندسة الزراعية، جامعة دمشق، سوريا.

الملخص

الأهداف: تهدف الدراسة إلى اقتراح تركيبة محصولية مثلى تزيد من صافي عائد وحدة المساحة في محافظة السويداء، سوريا.

المنهجية: باستخدام البرمجة الخطية الرياضية التي تعد أداة مهمة لدراسة الجوانب المختلفة للتراكيب المحصولية مع كل ما يواجهها من قيود تعيق الإنتاج، كالظروف الجوية المتغيرة ومشاكل المياه ومشكلات العمالة والظروف الاقتصادية. من خلال استبيان استهدف مزارعي الخضروات المروية الصيفية والشتوية خلال موسم 2020. لعينة مكونة من 106 مزارع، من خلال إعادة توزيع المساحة المخصصة لكل محصول.

النتائج: وجدت الدراسة زيادة بنسبة 80.8٪ في إجمالي الدخل الصافي مقارنة بنمط المحصول الفعلي مع تقليل استهلاك المياه.حيث بلغ حوالي 5.6 مليون م3 من المياه مقابل إجمالي استهلاك التركيبة المحصولية الفعلية لعينة المزارعين 5.9 مليون م3 من المياه.

الخلاصة: تشمل مساهمات الورقة اقتراح تركيب محصولي أمثل، وزيادة صافي الدخل، وتقليل استهلاك المياه، والتأكيد على أهمية استخدام البرمجة الخطية في صنع القرار الزراعي، مع الحفاظ على المساحة المخصصة لزراعة المحاصيل الأساسية كالقمح والبطاطا والبندورة والبطيخ.

الكلمات الدالة: التركيبة المحصولية، البرمجة الرياضية الخطية، تعظيم، الأمثل، محافظة السويداء.

Introduction

The Crop pattern is a relative concept because it is sometimes not easy to achieve what is precisely optimal. Where there are many obstacles, including: Limited arable land, sometimes the expansion of one crop excludes other crops, in addition to climatic factors and the risk factor in agriculture.

However, sometimes it is necessary to recombine the crop distribution for the scarcity of a resource or in response to some agricultural policies, where the optimal crop pattern can reflect the importance of allocation and planning in the use of the agricultural land to achieve the largest amount of production and reach the highest net income possible farm. Given the practical scenario, decision makers are facing challenges in terms of input management, the type of crop to be grown, and the different agricultural techniques applied to increase the farm production and return. These challenges are associated with socio economic development and the scarcity of resources in particular region.

To overcome these problems, a linear programming technique is applied in order to optimize the farm's returns by allocating the available farm resources optimally (Bhatia, 2019).

It is a mathematical modeling technique used to assist managers in planning and decision-making in order to optimally allocate limited resources, and it can be used to solve a wide range of emerging problems in business, agricultural, governmental, industrial, and others. It has proven its ability in various fields such as production, finance, marketing, research and development, and laborers management by determining the optimal product mix.

2. Literature Review

There are many studies, which have used the mathematical linear programming methodology in order to plan and optimally allocate the crop composition according to the available land and crop resources in order to achieve the best possible returns, including:The results of Mellaku et al. (2018), from a single rural community in Southern Ethiopia suggest that the overall performance of linear programming-based cropland allocation offer significantly higher returns than current cropland allocation methods. Under the status quo cropland allocation, households were not even able to meet the minimum required food crop production levels for all crops, while through a linearly programmed land use allocation the profit potential was more than doubled.

In the study of Osama et al. (2017), a linear optimization model developed to maximize the net annual return from the three old regions of Egypt. Data for 28 crops in five years from 2008 to 2012 was being analyzed. The spatial variations of crops, irrigation water needs, crop yields and food requirements incorporated in the model. The results show that there is a significant reduction in the allocated areas for onion, garlic, barley, flax, fenugreek, chickpeas, lentil and lupine since they considered as non-strategic crops. The allocated areas for strategic crops such as wheat, maize, clover, rice, sugar products and cotton remained almost the same to satisfy their actual food requirements. Crops with high net returns such as tomatoes have increased substantially. The developed model proposes a change in the cropping pattern in the old lands of Egypt to increase the gross net return without adding further any other expenses.

To solve a maximizing problem of gross margin among a group of farm data on a representative sample of farmers involved in arable crop farming in combination with monogastric farm animals and fish farming in Ohafia Stat/ Nigeria. Data were analyzed using linear programming, the proposed model recommended yam 0.29 ha, cassava 0.02 ha and cassava/maize/cocoyam 0.13 ha, broiler I – August –December 70 birds, 220 fish and layers 205 birds enterprises for an average farmer in Ohafia to optimize gross margin given the available resources. The optimum gross margin was 72.90% greater than obtained in the existing plan (Igwe et al., 2013).

Farmers have to choose among a variety of complex ways of production. Crop planning may involve choices about varieties, planting dates, fertilizer, and pesticide treatments. Linear programming has proved a very flexible tool for modeling these kinds of complexities and it was developed to determine the optimal crop pattern for a rural farmer. Crops considered: maize, soya beans and cotton. The proposed model produced an optimal crop combination that gives higher income than that obtained from the farmer's plan. The income difference was 72.79 percent. (Majeke et al., 2013).

A Linear Programming, to 2004-2005 data related to classic index (DRC) in Kermanshah province was applied to investigate comparative advantages of corn in comparison with other competitor crops. Results showed that comparing optimal cultivation pattern resulted from linear programming models with crops ranking based on comparative advantage indices indicated that, resources availability and limitations, tradable and non-tradable inputs costs and yield would lead to shift in production comparative advantage from one crop to another. Factors such as supporting policies and rotation might also have effects on comparative advantages and optimal cultivating pattern (Abedi et al., 2011).

Previous reference studies have shown the importance of adopting linear programming as an important tool for managing the various aspects of crop pattern, with all the constraints facing that hold back achieving the wanted production goal. Therefore, this study came to show how to find a solution to the problems of land use planning in the targeted areas in

As-Swaida Governorate, in order to reach the optimal production of the Irrigated crops pattern. By formulating a model in which the crop pattern is restructured and the crop areas and productive resources are redistributed, which contributes to the sustainability of agricultural resources, based on the importance of this sector at the governorate level in order to achieve self-sufficiency on the one hand, and the need to statement the study of this aspect of optimal resource management. On the other hand.

The low efficiency of using the available arable land for the current irrigated crop pattern is one of the most important risks that is facing agriculture in As-Swaida Governorate. Where farmers usually tend to produce high-profit crops without considering the balance in the distribution of other crops, and this leads to an uneconomic depletion and degradation of the most important resource which is the land.

The importanceof this research lies in the possibility of presenting different alternatives of the agricultural cropping pattern in the governorate that will benefit decision makers, and maximize the return within the available land resources.

The main objective of this research is to maximize the net return of the irrigated vegetable area in Swaida Governorate by applying the mathematical linear programming method, by achieving the following:

- Determining the components of the actual irrigated agricultural crop pattern.
- Reaching the best agricultural crop pattern that maximizes the net return per unit area within the constraints of the available land resources.

3. Materials and methods

3.1 Data and study area: The research was based on:

- Primary data: through a questionnaire targeting the farmers of irrigated crops. Moreover, it included questions related to the composition of crops approved by farmers especially with regard to crop water needs, production costs, and current agricultural prices in As-Swaida Governorate / southern Syria during 2020 season.

The average total irrigated land during the period between 2016 -2020 was about 3481 hectares, of which the percentage of irrigated land from wells was about 59.9%. and irrigated from government irrigation projects amounted to about 40.1%, and in contrast the percentage of irrigated land according to modern irrigation methods (Mist irrigation, drip irrigation) about 94% of the average total irrigated land in As-Swaida during the period between 2016 -2020 (statistics of the Ministry of Agriculture, 2016-2020).

- Secondary data: were also relied on the number of irrigation wells that have been operating for at least three consecutive years to irrigate vegetable crops. Where the target community was identified as owners of these wells, which amounted to (221) wells (Department of Agricultural Extension, 2020),

3.2 The sample size: according to the following law: (Glenn, 1992) (Yamane, 1967)

$$n = \frac{N}{1 + N(e)^2}$$

Where:

N: The studied community (221) wells.

e: Precision Level %7± Where Confidence Level is 95%.

n: The sample size is consisted of (106) observations that represented 48.06% of the studied statistical community.

3.3 Statistical Methods: The study relied on several methodologies in economic analysis:

- Descriptive analysis methods: to describe the study variables such as means, percentage
- Quantitative analysis in management: by using of one of the models of operations research is a Linear Programming Model Which is used for resource allocation when the resources are limited and there are a number of competing candidates for the use of resources and may be used to maximize the returns or minimize the costs (Murthy, 2007).

The construction of the mathematical model for any problem goes through the following: Formulation of the problem: Optimization problems are often formulated in a verbal form, and the method of solution is determined by depicting the problem in the form of a model of a mathematical program, and then solving this program according to the known methods of solution, through the following entrance:

1. Determine the quantities that need optimal values, and express them with mathematical functions. This procedure helps to determine the inputs and outputs.
2. define the demands, restrictions, and limits, and express them mathematically, and these demands constitute the imposed restrictions.

3. Express any conditions other than apparent, which include the absence of a negative value, or adherence to integers in the input of the variables (Bronson, 1988).

The Linear Programming Model of the optimal cropping pattern that achieves the maximum total return from an area unit takes the following form:
Max: P = p1X1 + p2X2 + p3X3+……+ pnXn
S/C:
a11 X1+ a12 x2+ a13 X3+ …… + a1n Xn ≤ b1
a21 X1+ a22 X2+ a23 X3+ …… + a2n Xn ≤ b2
a31 X1+ a32 X2+ a33 X3+ …… + a3n Xn ≤ b3
am1 X1+ am2 X2+ am3 X3+ …… + amn Xn ≤ bm
X1 ≥ 0, X2 ≥ 0, X3 ≥ 0, ……,Xn ≥ 0

Where:
Max P: The objective function in Maximization, which is the sum of the net returns from different crops of the crop pattern, (Syrian Pound).
X1…Xn: Decision variables of the linear program to be defined, representing the areas of the crops constituting the crop pattern (Dunums).
P1…P2: the variables' parameters that effecting on the function, representing the net yield of a dunum of the different crops (S.P/dunum).
a11…a1n: The parameters of a linear program's variable used by each resource and are known.
b1…bm: represents the available resources that are known and specific.
And all xj are ≥ 0: are non-negativity constraint.
These programs were used to analyze and process data:
IBM SPSS 28: for descriptive and quantitative data analysis.

Excel Solver: for solving optimization problems in linear mathematical programming.

4. Results and Discussion

4.1 Actual Components of the Crop pattern

Included 12 summer crops represented: tomatoes, eggplants and red watermelons the three highest percentages were 81%, 39%, and 35%, respectively of the sample farmers. In addition, 7 winter crops represented the three highest: Cauliflower, Peas and wheat, with rates amounting to 21%, 21%, and 16% of the sample farmers, figures (1) and (2) show the rest of the crops of the crop composition.

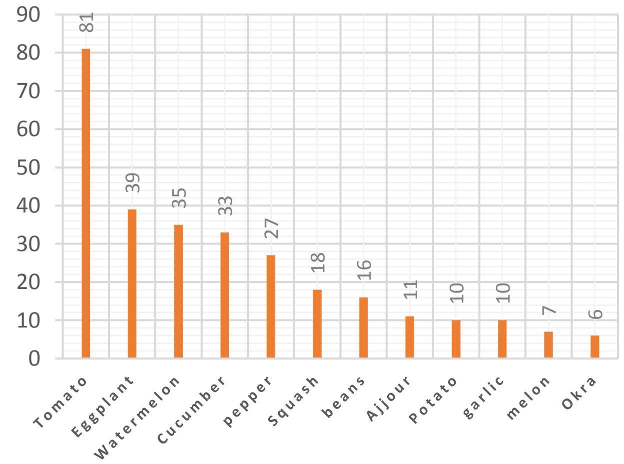

Figure 1. The percentage of summer crops in the actual crop irrigated pattern.

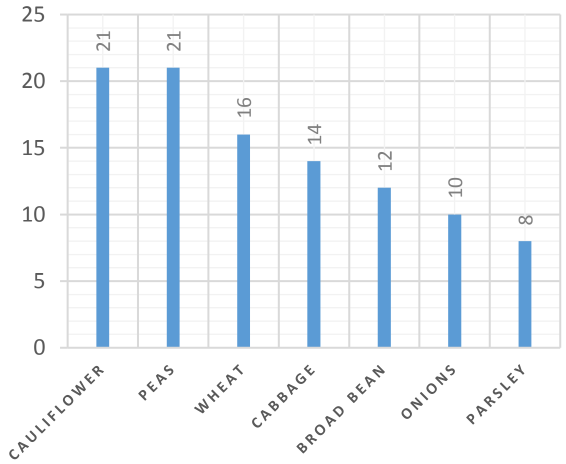

Figure 2. The percentage of winter crops in the actual crop irrigated pattern.

4.2 Actual Cultivated Irrigated Area

Table (1) shows the distribution of the actual cultivated, the study sample varied in terms of the actual cultivated area, between a minimum of one Dunum, and a maximum of 310 D, with an average of about 77.55 D in studying sample. The summer Cultivated area varied from between a minimum of one Dunum, and a maximum of 285 Dunums and an average of about 61.9 Dunums and a total area of 6,065.9 Dunums. In contrast, the total area of winter formula cultivated by sample farmers was 2050.4 Dunums, with an average of about 35.8 Dunums.

Table 1. Minimum and maximum limits of the area of the actual cultivated irrigated areas for the 2020 season. Dunum

	Min	Max	Sum	Mean	Std. Deviation
Total Actual Cultivated Irrigated Area	1.00	310.00	8116.3	77.55	79. 9
Total summer Cultivated area	1.00	285.00	6065.90	61.9	62.6
Total winter Cultivated area	.50	165.00	2050.40	35.8	39.85

Source: Questionnaire.

4.3 The optimal cropping pattern that maximizes the net return per unit area

This paragraph includes the analysis of linear programming model for the optimal irrigated crop pattern that maximizes the net return in the study area, where an analytical framework was built to reach the optimal cropping pattern by suggesting: alternative model: An optimum cropping pattern that maximizes the net return per unit area within the irrigated area that actually cultivated.

4.3.1 Mathematical form

- **Objective function:**

$$\text{MAX: } 1936.4 * X1 + 207911 * X2 + 350619 * X3 + 1928 * X4 + 101921 * X5 + 29939.2 * X6 + 159908 * X7 + 22687.2 * X8 + 32526.2 * X9 + 13209.5 * X10 + 78421.2 * X11 + 101728 * X12 + 1.89305E+006 * X13 + 16812.5 * X14 + 45562.5 * X15 + 38966.7 * X16 + 1254 * X17 + 149082 * X18 + 12879.1 * X19$$

Max Z: Maximizing the net return per area unit for all crops.
Ci: net return per area unit for each crop.
Xi: Crops.

- **Constraints:**

a. **Land Constraints: which includes three constraints:**

- The total area of irrigated winter crops equal or less than the total irrigated winter area.

$$X13 + X14 + X15 + X16 + X17 + X18 + X19 <= 2050.4$$

- The total area of irrigated summer crops equal or less than the total irrigated summer area.

$$X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8 + X9 + X10 + X11 + X12 <= 6065.9$$

- Total irrigated area equals the total of the winter and summer area.

$$X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8 + X9 + X10 + X11 + X12 + X13 + X14 + X15 + X16 + X17 + X18 + X19 = 8116.3$$

b. **Structural constraints: The minimum and upper limits have been set for the area of each crop:**

- The minimum limit for each crop is equal or greater than 10% of the total area of the crop, except (tomatoes, melons, potatoes and wheat not less than 50%).

Minimum limit tomato	X1 >= 1770.7
Min. lim. eggplant	X2 >= 27.16
Min. lim. pepper	X3 >= 16.05
Min. lim. watermelon	X4 >= 119.45
Min. lim. melon	X5 >= 10.05
Min. lim. cucumber	X6 >= 19.6

Min. lim. squash	X7 >= 7.74
Min. lim. Armenian cucumber	X8 >= 4.6
Min. lim. okra	X9 >= 1.6
Min. lim. beans	X10 >= 4.5
Min. lim. potato	X11 >= 31.7
Min. lim. onion	X12 >= 10
Min. lim. garlic	X13 >= 2.8
Min. lim. cabbage	X14 >= 16.01
Min. lim. cauliflower	X15 >= 35.31
Min. lim. peas	X16 >= 32.59
Min. lim. Broad Beans	X17 >= 14.99
Min. lim. parsley	X18 >= 2.35
Min. lim. wheat	X19 >= 504.95

- The upper limit of each crop does not exceed the total area of the cultivated crop, except (tomatoes, wheat, potatoes, and red watermelons).

maximum limit eggplant	X2 <= 271.6
Max. lim. pepper	X3 <= 160.5
Max. lim. melon	X5 <= 100.5
Max. lim. cucumber	X6 <= 196
Max. lim. squash	X7 <= 77.4
Max. lim. Armenian cucumber	X8 <= 46
Max. lim. okra	X9 <= 16
Max. lim. beans	X10 <= 45
Max. lim. onion	X12 <= 100
Max. lim. garlic	X13 <= 28
Max. lim. cabbage	X14 <= 160.1
Max. lim. cauliflower	X15 <= 353.1
Max. lim. peas	X16 <= 325.9
Max. lim. Broad Beans	X17 <= 149.9
Max. lim. parsley	X18 <= 23.5

a. **non negativity constraint: $\mathbf{x_{ij} \geq 0}$**

4.3.2 Crops area of the proposed pattern

The results of solving the programming model showed that to achieve greater net return per unit area, it required increasing the cultivated area of some crops, decreasing others, and stabilizing others as follows:

Increasing the area of: potato and wheat crops was estimated at about 3118.45, 134.91 Dunums, respectively, at a rate of 983.7%, 13.35% over the actual, respectively.

Reducing the area of: okra, beans, Armenian cucumber, beans, cucumber, watermelon, tomato by about 14.4, 40.5, 41.4, 134.91, 176.4, 1075.05, 1770.7 Dunum, in order. Unchanged of the rest of crops. Table (2).

Table 2. The Crops area of the proposed pattern according to the programming model compared to the actual cropping pattern.

Code	Crop	Actual area/Dunum	Proposal Area/ Dunum	The difference between the actual and the proposed / acres	The difference ratio/%
x11	Potato	317	3435.45	3118.45	983.73
x20	Wheat	1009.9	1144.81	134.91	13.36
x2	Eggplant	271.6	271.6	0	0
x3	Pepper	160.5	160.5	0	0
x5	Melon	100.5	100.5	0	0
x7	Squash	77.4	77.4	0	0
x12	Onions	100	100	0	0
x13	Garlic	28	28	0	0
x14	Cabbage	160.1	160.1	0	0
x15	Cauliflower	353.1	353.1	0	0
x16	Peas	325.9	325.9	0	0
x19	Parsley	23.5	23.5	0	0
x9	Okra	16	1.6	-14.4	-90
x10	Beans	45	4.5	-40.5	-90
x8	Armenian cucumber	46	4.6	-41.4	-90
x17	Broad Beans	149.9	14.99	-134.91	-90
x6	Cucumber	196	19.6	-176.4	-90
x4	Watermelon	1194.5	119.45	-1075.05	-90
x1	Tomato	3541.4	1770.7	-1770.7	-50
Total		8116.3	8116.3	-	-

Source: The results of Solver Program.

4.3.3The net return of the proposed cropping pattern

The proposed model achieved an increase in the total net return about 80.8% over the net return of the actual pattern, where achieved 522.16 million SP, compared to 288.76 million SP of the actual pattern. Table (3) shows the changes in the net return that achieved from the proposed cropping pattern for each crop whose area has increased, decreased, or remained constant.

4.3.4 Comparison between the proposed scenario and the actual cropping pattern

The proposed scenario did not exclude any of the crops, and it achieved a positive effect on the net return and the quantity of irrigation water, Table 4, figures (3) and (4) where: Scenario exceeded the total net return per unit of area about 522.15 million SP, compared to 288.76 million SP, for the total net return realized from the area unit for the actual cropping.

Scenario exceeded the total amount of water consumed, with the lowest consumption of irrigation water, amounting to approximately 5.6 million m3 of water, while the total consumption of the actual cropping by sample farmers was 5.9 million m3 of water.

Table 3. The total net return of the proposed pattern.

Code	Crop	Actual Return per D / SP * 1000	Total Return for all area/ SP *million		The difference/ SP *million	The difference ratio/%
			proposed	Actual		
x11	Potato	78.42	269.41	24.86	244.55	983.74
x20	Wheat	12.88	14.74	13.01	1.74	13.36
x2	Eggplant	207.91	56.47	56.47	0	0
x3	Pepper	350.62	56.27	56.27	0	0
x5	Melon	101.92	10.24	10.24	0	0

Code	Crop	Actual Return per D / SP * 1000	Total Return for all area/ SP *million		The difference/ SP *million	The difference ratio/%
			proposed	Actual		
x7	Squash	159.91	12.38	12.38	0	0
x12	Onion	101.73	10.17	10.17	0	0
x13	Garlic	1893.05	53.01	53.01	0	0
x14	Cabbage	16.81	2.69	2.69	0	0
x15	Cauliflower	45.56	16.09	16.09	0	0
x16	Peas	38.97	12.70	12.70	0	0
x19	Parsley	149.08	3.50	3.50	0	0
x17	Broad Beans	1.25	0.02	0.19	-0.17	-90
x9	Okra	32.53	0.05	0.52	-0.47	-90
x10	Beans	13.21	0.06	0.59	-0.53	-90
x8	Armenian cucumber	22.69	0.10	1.04	-0.94	-90
x4	Watermelon	1.93	0.23	2.30	-2.07	-90
x1	Tomato	1.94	3.43	6.86	-3.43	-50
x6	Cucumber	29.94	0.59	5.87	-5.28	-90
	Total		522.16	288.76	233.40	80.83

Source: The results of Solver Program and table (2).

Table 4. Efficiency of the proposed scenario is compared to the actual cropping pattern.

	Total area / Dunum	total return per unit of area / million sp.	total optimum water requirement / million m^3
actual cropping	8116.3	288.76	5.90
the proposed scenario	8116.3	522.15	5.63

Source: Calculated and collected based on Appendix 1 data.

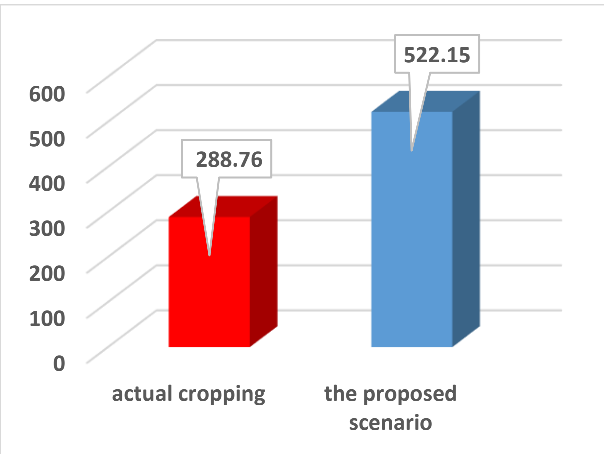

Figure 3. The total return per unit of area / million sp.

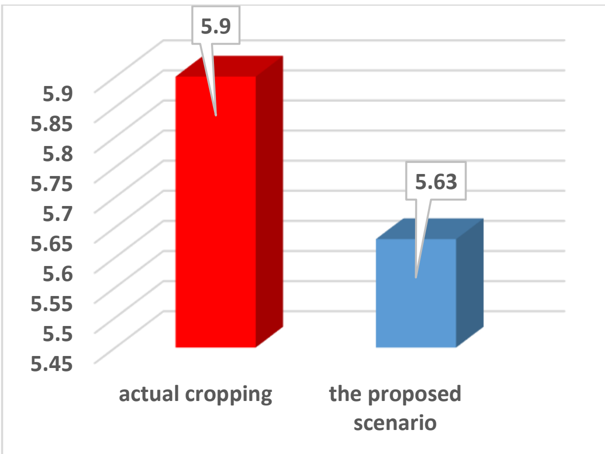

Figure 4. The total optimum water requirement / million m3

5. CONCLUSION

The aim of the study is to explore and characterize model of crop pattern that lead to maximizing the net return. The study applied the linear programming method to achieve its objective: An optimum cropping pattern that maximizes the net return per unit area within the irrigated area that actually cultivated. The model achieved an increase in the total net return about 80.8% over the net return of the actual pattern, where achieved 522.16 million SP, compared to 288.76 million SP of the actual pattern. And it was reduced the total amount of water consumed, with the lowest consumption of irrigation water, amounting to approximately 5.6 million m3 of water, while the total consumption of the actual cropping by sample farmers was 5.9 million m3 of water.

Notes and recommendations:

- The increase in total revenue by 80.8% in the proposed composition resulting from the redistribution of the area allocated to each crop shows the effectiveness of using programming models as one of the decision-making methods that help decision makers in managing available resources, especially (land, water)
- Modifying the cropping pattern in Swaida Governorate to achieve the optimal cropping pattern that maximizes the net return from the unit area of agricultural activities, while preserving the area allocated for the cultivation of basic crops wheat, tomatoes, potatoes and watermelon, within the limits of the available land resources, that achieve self-sufficiency in the governorate.
- Through the lower and upper limits of the crop areas in the constraints of the proposed models, the flexibility and effectiveness of using mathematical linear programming in the management of agricultural resources is shown.They are subject to change according to the conditions imposed by the production process or according to the strategic criteria desired by the decision-maker, with the possibility of taking into account environmental conditions and preserving strategic crop areas in order to achieve self-sufficiency.
- The method of constructing a mathematical linear programming model from an objective function and constraints makes each constraint a partial goal that must be achieved to reach the general goal of the model. For example: The minimum area constraint for main crops such as tomatoes, watermelon, potatoes, and wheat is not less than 50% of the actual area in the composition, as it is a binding constraint to achieve the general goal of the total return maximization model. Thus, the constraints of the proposed model could include many goals related to sustainability, agricultural labor or agricultural mechanization, to serve the case of multiple goals and decisions facing the decision-maker.

6. REFERENCES

- Abedi, S., Peykani, G.R., and Kalashami, M.K. (2011). Determining Comparative Advantages of Corn in Optimal Cultivation Pattern. *International Journal of Agricultural Management & Development*. 1(4):197-206.
- Bhatia, M. 2019. Linear Programming Approach- Application in Agriculture. *Journal of Emerging Technologies and Innovative Research*. 6(5):155-157.
- Bronson, R. (1988). *Schaum's outline of theory and Problems of Operations Research*, Translated by: Hassan Hosni Al-Ghobari. International House for Publishing and Distribution. Cairo.
- Glenn, D.I. (1992). *Determining Sample Size*. Florida Cooperative Extension Service. University Of Florida. Fact Sheet Peod-6.
- Igwe, K.C., Onyenweaku, C.E., and Tanko, L. (2013). A Linear Programming Approach To Combination Of Crop, Monogastric Farm Animal And Fish Enterprises In Ohafia Agricultural Zone, Abia State, Nigeria. *Global Journals Inc. (US).*13(3):23-32.
- Majeke, F., Mubvuma, M.T., Makaza, K., and Mutambara, J. (2013). Optimum Combination of Crop Farm Activities: Application of a Linear Programming Model to a Rural Farmer in Zimbabwe. *Greener Journal of Economics and Accountancy.* 2(2):58-61.
- Mellaku, M.T., Reynodlds, T.W., and Woldemamanual, T. (2018). Linear Programming-Based Cropland Allocation to Enhance Performance of Smallholder Crop Production: A Pilot Study in Abaro Kebele, Ethiopia. *Resources.* 7(76):2-15, https://doi.org/10.3390/resources7040076

- Ministry of Agriculture and Agrarian Reform. Directorate of Agricultural Extension unpublished Data 2020. Swaida, Syria.
- Murthy, P. Ram. (2007). *Operations research (2nd ed.)*. New Age International (P) Ltd., Publishers.
- Osama, S., Elkholy, M., Kansoh R. M. (2017). Optimization of the cropping pattern in Egypt. *Alexandria Engineering Journal*. 56, 557–566
- Yamane, T. (1967). *An Introductory Analysis* 2. Nd Ed., New York: Harper and Row.

Appendix 1

Crop	Code	The optimum water requirement m^3 /Dunum	Actual Return per D / SP * 1000	AREA / Dunum		Total Optimum Water Requirement m^3 /Dunum		Total Return / SP * 1000	
				Proposal	Actual	Proposal	Actual	Proposal	Actual
Tomato	X1	897.9	1.94	1770.7	3541.4	1589911.5	3179823.1	3435.158	6870.316
Eggplant	X2	782.6	207.91	271.6	271.6	212554.16	212554.16	56468.36	56468.36
Pepper	X3	779.1	350.62	160.5	160.5	125045.55	125045.55	56274.51	56274.51
Watermelon	X4	698.2	1.93	119.45	1194.5	83399.99	833999.9	230.5385	2305.385
Melon	X5	698.2	101.92	100.5	100.5	70169.1	70169.1	10242.96	10242.96
Cucumber	X6	486.3	29.94	19.6	196	9531.48	95314.8	586.824	5868.24
Squash	X7	505.9	159.91	77.4	77.4	39156.66	39156.66	12377.03	12377.03
Armenian cucumber	X8	584.6	22.69	4.6	46	2689.16	26891.6	104.374	1043.74
Okra	X9	747.6	32.53	1.6	16	1196.16	11961.6	52.048	520.48
Beans	X10	362.7	13.21	4.5	45	1632.15	16321.5	59.445	594.45
Potato	X11	685.9	78.42	3435.45	317	2356375.2	217430.3	269408	24859.14
Onions	X12	992	101.73	100	100	99200	99200	10173	10173
Garlic	X13	182.2	1893.05	28	28	5101.6	5101.6	53005.4	53005.4
Cabbage	X14	554	16.81	160.1	160.1	88695.4	88695.4	2691.281	2691.281
Cauliflower	X15	554	45.56	353.1	353.1	195617.4	195617.4	16087.24	16087.24
Peas	X16	73	38.97	325.9	325.9	23790.7	23790.7	12700.32	12700.32
Broad Beans	X17	129.9	1.25	14.99	149.9	1947.201	19472.01	18.7375	187.375
Parsley	X19	410.5	149.08	23.5	23.5	9646.75	9646.75	3503.38	3503.38
Wheat	X20	623	12.88	1144.81	1009.9	713216.63	629167.7	14745.15	13007.51
Total				8116.3	8116.3	5628876.8	5899359.8	522163.75	288780.12

Source: Questionnaire.