These signals can be stored on thinkspeak server or on any other server [11] (shown by Laptop in block diagram in figure 2).

Conclusion. We can conclude that this paper focuses on design of a system in which all the devices, gadgets, appliances that operates from Brain. It will remove all hurdles which are used in previous systems. That means we don't need to use any intermediate devices to control any other devices, appliances etc.

Acknowledgments. The report was published with the financial support of Savitribai Phule Pune University (formaly known as Pune University) under ASPIRE scheme.

References

1. Lee, M.-H. Development of an open source platform for brain-machine interface: openBMI / M.-H. Lee, S. Fazli, K.-T. Kim, S.-W. Lee // 2016 4th International Winter Conference on Brain-Computer Interface (BCI). – 2016.
2. Using Brain-Computer Interface and Internet of Things to Improve Healthcare for Wheelchair Users // UBICOMM 2017: The Eleventh International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies. – 2017. – P. 112–118.
3. Internet of Things Meets Brain–Computer Interface: A Unified Deep Learning Framework for Enabling Human-Thing Cognitive Interactivity // IEEE Internet of Things Journal. – 2019. – Vol. 6, № 2. – P. 2104–2114.
4. Boucha, D. Controlling electronic devices remotely by voice and brain waves / D. Boucha, A. Amiri, D. Chogueur // 2017 International Conference on Mathematics and Information Technology (ICMIT). – 2017.
5. Yadav, P. Application of Internet of Things and Big Data towards a Smart City / P. Yadav, S. Vishwakarma // 2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU). – 2018. – P. 1–6.
6. Real Time Emotion Detection From Speech Using Raspberry Pi-3 / A. Mishra, D. Patil, N. Karkhanis [et al.] // 2017 IEEE International Conference on Computing, Communication and Automation (ICCCA). – 2017. – P. 1342–1346.
7. Using Brain Waves to Control Computers and Machines / C. Braun, D. Pantazis, S. R. Soekadar, P. Bamidis // Computational Intelligence and Neuroscience. – 2013. – Vol. 2013. – Art. 802063. – DOI: 10.1155/2013/802063.
8. Yordanov, Y. Humanoid robot control with EEG brainwaves / Y. Yordanov, G. Tsenov, V. Mladenov // 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). – 2017. – P. 456–461.
9. Abdulkader, S. N. Brain computer interfacing: Applications and challenges / S. N. Abdulkader, A. Atia, M. Sami // Egyptian Informatics Journal. – 2015. – Vol. 16, № 2. – P. 213–230.
10. Beigh, N. T. A Review on Brain Wave Signal Appliance Control / N. T. Beigh, F. T. Beigh // International Journal for Research in Applied Science & Engineering Technology (IJRASET). – 2018. – Vol. 6, № 4. – P. 1892–1898.
11. Patil, D. P. IoT-Based Automatic Corona Virus Detection and Monitoring System / D. P. Patil, A. Mishra, T. Jaware // Health Informatics and Technological Solutions for Coronavirus (COVID-19). – Taylor & Francis, 2021. – P. 267–285.
12. Health Detection System for COVID-19 Patients Using IoT / D. P. Patil [et al.] // Medical Imaging and Health Informatics. – Wiley, 2022. – P. 189–205.

UDC 681

DETECTING EMAIL SPAM OR HAM USING LOGISTIC REGRESSION AND STREMLIT DEPOLYMENT

Pramod S. Aswale[1], Dr. Dipak P Patil[2]

[1]Savitribai Phule Pune University Pune, India
[2]Research Scholar, Sandip Institute of Technology and Research Centre, Nashik, India

Abstract. The use of email as a key form of communication has led to an increase in uninvited and potentially hazardous spam messages. This study presents a comprehensive approach to classifying emails as either spam or ham using logistic regression. Leveraging a dataset obtained from Kaggle, we performed extensive exploratory data analysis (EDA) with visualizations such as pie charts, bar plots, and confusion matrices. Using performance parameters like accuracy, precision, recall, and F1 score, the model's potential for prediction was assessed. In order to facilitate smooth interaction for real-time email classification, Streamlit was also used to develop an intuitive application. This work demonstrates the utility of logistic regression in tackling the spam problem while paving the way for further exploration using advanced machine learning models.

Keywords: EDA, ML, F1score.

Address for correspondence: Sandip Institute of Engineering and Management, Nashik, Maharastra, India
e-mail: ar.chaudhari1@gmail.com

Introduction. Email communication is integral to modern digital interactions, but it is increasingly plagued by spam messages that waste time, reduce productivity, and pose security risks. Spam filters are vital for distinguishing between spam and legitimate (ham) emails. Machine learning (ML) offers automated, scalable solutions to this problem. The usefulness of «logistic regression», a popular supervised machine learning algorithm, for spam detection is investigated in this work. This study has three goals in mind:

To use exploratory data analysis in order to comprehend and visualise the dataset.

To build and evaluate a logistic regression model for spam classification.

To deploy the model using Streamlit for real-world usability.

Dataset Description.

The dataset, sourced from Kaggle, contains thousands of labeled emails. Each email is classified as either spam or ham, with features such as email content, metadata, and sender information[1].

– Dataset Characteristics:

– Total Records: 5,572 emails.

– Features: “Text” (email content), “Label” (spam or ham).

To ensure a reliable model evaluation, the dataset was divided into training (80 %) and testing (20 %) subsets. Tokenisation, stop-word removal, and TF-IDF (“Term Frequency-Inverse Document Frequency”) feature extraction were among the text preparation methods used.

Methodology.

Data Preprocessing.

Spam emails a menudo contian noise tales como HTML tags, punctuations charácteres especiales.processing involves:

– Converting text to lowercase.

– Removing HTML tags and non-alphanumeric characters.

– Tokenizing and stemming.

– Extracting features using TF-IDF to convert textual data into numerical vectors for model training.

Model Training

Because of its ease of use and efficiency in binary classification tasks, logistic regression was chosen. The preprocessed training dataset was used to train the model. Grid search was used for hyperparameter optimization in order to maximize performance.

Deployment

The trained model was integrated into a Streamlit application, providing users with an intuitive interface for email classification. Users can input email text into the app and instantly receive predictions indicating whether the email is spam or ham.

Deployment. The trained model was integrated into a Streamlit application, providing users with an intuitive interface for email classification. Users can input email text into the app and instantly receive predictions indicating whether the email is spam or ham.

Exploratory Data Analysis (EDA).

Visualizations were employed to gain insights into the dataset:

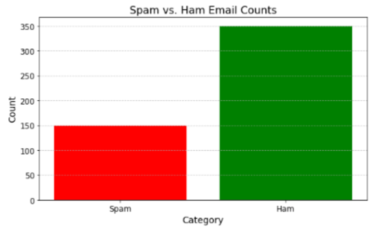

Figure 1 – The frequency of key terms indicative of spam, such as “free”, “win”, and “money”

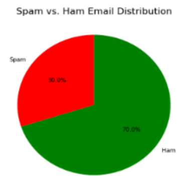

Figure 2 – Distribution of Spam versus Ham Emails

The figure 2 revealed a clear imbalance between spam and ham emails, emphasizing the importance of addressing class imbalance during model training. Key terms identified in bar plots provided insight into patterns that distinguish spam emails.

Results and Discussion.

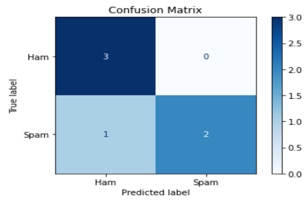

Figure 3 – Confusion Matrix: Evaluate model Performance

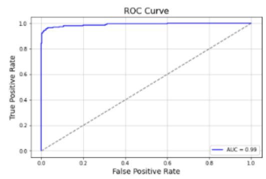

Figure 4 – Performance Evaluation using ROC curve

Table 1 – Model Performance

SStudy/Paper	Method / Algorithm	Accuracy	Precision	Recall	F1 Score	Remark
This Study	Logistic Regression	98 %	94 %	92 %	93 %	Balanced precision and recall; Streamlit deployment for real-time use
Ekramul Haque Tusher et al., 2024	Enhanced Logistic Regression	Up to 98 %	Not specified	Not specified	Not specified	Feature engineering with TF-IDF, RFE, PCA
Kaggle Dataset Example, 2025	Logistic Regression	99 %	Not specified	Not specified	Not specified	High accuracy on Kaggle dataset
Asif Karim et al., 2019	Various Machine Learning	Up to 99 %	Not specified	Not specified	Not specified	Includes logistic regression among best models

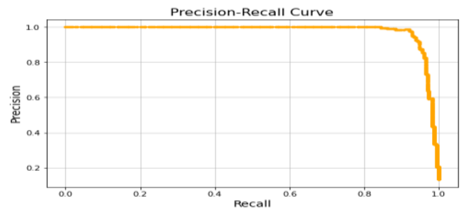

Figure 5 – Performance Evaluation using ROC curve

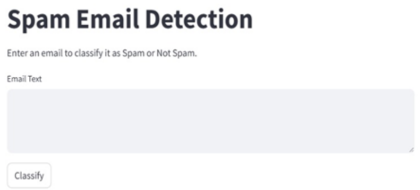

Figure 6 – Deployment Screen Showcasing Funtionality

The confusion matrix revealed minimal false positives and false negatives, demonstrating the model's reliability.

Deployment Results. The Streamlit application performed well in real-world scenarios, offering immediate and accurate predictions. The interactive interface allowed users to experiment with different email inputs and understand the model's behavior.

Conclusion. This study demonstrates how well logistic regression works to detect email spam. Performance measurements and visualizations highlight the model's resilience. Real-time email classification is made possible by the Streamlit application, which increases usefulness. Even while logistic regression worked well, more sophisticated machine learning models like random forests or deep learning might be investigated in future studies to further improve accuracy.

References

1. Kaggle. Email Spam Detection Dataset. – 2025. – URL: https://www.kaggle.com/datasets/ (date of access: 15.11.2024).

2. Pedregosa, F. Scikit-learn: Machine Learning in Python / F. Pedregosa, G. Varoquaux, A. Gramfort [et al.] // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2825–2830.

3. Streamlit. Documentation Streamlit. – 2025. – URL: https://docs.streamlit.io/ (date of access: 15.11.2024).

4. Han, J. Data Mining: Concepts and Techniques / J. Han, M. Kamber, J. Pei. – 3rd ed. – Waltham: Morgan Kaufmann, 2011. – 703 p.

5. Haque, E. Email Spam: A Comprehensive Review of Optimized Detection Methods, Challenges, and Open Research Problems / E. Haque, M. A. Ismail, M. A. Rahman, A. H. Alenezi, M. Uddin // IEEE Access. – 2024. – Vol. 12. – P. 45678–45699.

6. A Comprehensive Survey for Intelligent Spam Email Detection / A. Karim, S. Azam, B. Shanmugam, K. Kannoorpatti, M. Alazab // IEEE Access. – 2019. – Vol. 7. – P. 168261–168295.

УДК 621.3.049.76

ЛАЗЕРНАЯ ОБРАБОТКА КРЕМНИЯ ДЛЯ СОЗДАНИЯ МИКРОМЕХАНИЧЕСКИХ ЧУВСТВИТЕЛЬНЫХ ЭЛЕМЕНТОВ

Реутская О. Г., Таратын И. А., Луданик Д. Ю., Койро А. В., Войтеховский А.А.

Белорусский национальный технический университет, Минск, Беларусь

Аннотация. В работе рассматривается использование лазерной обработки кремния для создания высокочувствительных датчиков угла наклона и газового микроанализа. Проведён анализ технологических особенностей формирования и её влияния на параметры сенсоров.

Ключевые слова: кремний, чувствительный элемент угла наклона и газового микроанализа, лазерная обработка, навигационные системы.

LASER PROCESSING OF SILICON TO CREATE MICROMECHANICAL SENSING ELEMENTS

Reutskaya O. G., Taratyn I. A., Ludanik D. Yu., Koiro A. V., Voitekhovsky A. A.

Belarusian National Technical University, Minsk, Belarus

Abstract. The work considers the use of laser treatment of silicon to create high-sensitivity angle sensors and gas microanalysis. The analysis of technological formation and its influence on sensor parameters was carried out.

Key words: silicon, tilt angle sensitive element and gas microanalysis, laser processing, navigation systems.

Адрес для переписки: Реутская О. Г., пр. Независимости, 65, г. Минск 220113, Республика Беларусь
e-mail: oreutskaya@gmail.com

Введение. Навигационные системы требуют высокой точности и стабильности при измерении углов наклона в условиях вибраций и температурных колебаний. Одним из перспективных направлений является использование лазерной обратки кремния при формировании конструкции чувствительного элемента. Датчик наклона измеряет изменение положения массы под действием силы тяжести. Внутри кремниевого чипа формируется подвижная механическая структура – инерционная масса, подвешенная на упругих подвесах толщина подвесов [1].

Методы и материалы. Датчики угла наклона, как правило, относятся к категории инерциальных микромеханических датчиков, куда также входят акселерометры и гироскопы. Их производство базируется на стандартных процессах микроэлектроники, адаптированных для создания трехмерных механических структур. В производстве МЭМС используется монокристаллический