SURVEY ON PASSWORD STRENGTH ANALYZER USING LSTM AND CNN

BINOY SHIJU,ARATHY KL,ANRIYA JAISON,ADHINADH MANOJ

Binoy Shiju,B.Tech Computer Science , Holy Grace Academy Of Engineering
Arathy K L,B.Tech Computer Science , Holy Grace Academy Of Engineering
Anriya Jaison,B.Tech Computer Science , Holy Grace Academy Of Engineering
Adhinadh Manoj,B.Tech Computer Science , Holy Grace Academy Of Engineering
Ajith P J,Mentor,Holy Grace Academy Of Engineering
Sanam E Anto, Head of the Department(Computer Science),Holy Grace Academy Of Engineering

ABSTRACT

The Password Strength Analyzer using AI and Machine Learning is a web-based application designed to assess and enhance password security using a hybrid LSTM-CNN model. The system comprises a user-friendly frontend developed with HTML, CSS, and JavaScript, which allows users to input their passwords and receive instant feedback on their strength, score, and recommendations for improvement. The backend, built with Flask and TensorFlow, uses a pre-trained deep learning model that integrates Conv1D (CNN) for local pattern extraction and LSTM layers for capturing sequential dependencies, ensuring precise classification into Weak, Medium, or Strong categories. The model is trained on a large password dataset, tokenized at the character level, and uses padded sequences to maintain consistency. Additionally, the system performs real-time leak detection by querying the "Have I Been Pwned" API, checking if the password has appeared in known data breaches. The application also offers actionable recommendations, such as adding uppercase letters, digits, or special characters, to help users create stronger, more secure passwords. This tool provides an effective and interactive solution for promoting better password practices and strengthening online security.

Keywords: - *LSTM(Long Short Term Memory Machine), CNN(Convolution Neural Network), Flask, Tensorflow, Conv1D*

1.Introduction

In today's digital landscape, where data breaches and cyberattacks are increasingly common, the importance of strong password security cannot be overstated. Weak or compromised passwords remain one of the leading causes of unauthorized access and data theft. To address this issue, the Password Strength Analyzer using AI offers an intelligent solution for evaluating and improving password security. This project leverages artificial intelligence (AI) and deep learning to provide accurate and reliable password strength assessment, helping users create robust credentials to safeguard their sensitive information.The system utilizes a hybrid LSTM-CNN model to analyze password complexity effectively. The CNN (Conv1D) extracts local patterns from password sequences, while the LSTM layers capture long-term dependencies, allowing the model to classify passwords into three categories: Weak, Medium, and Strong. To further enhance security, the application integrates a real-time leak detection feature using the "Have I Been Pwned" API, which checks if the password has been involved in known data breaches.The frontend interface, built with HTML, CSS, and JavaScript, offers a user-friendly experience by allowing users to input their passwords and receive instant feedback on strength, score, and improvement suggestions. The backend, powered by Flask and TensorFlow, processes the input and returns the results efficiently. The model is trained on a diverse dataset of passwords, tokenized at the character level, and padded for consistency, ensuring accurate predictions.By providing actionable recommendations, such as incorporating uppercase letters, digits, and special characters, the Password Strength Analyzer empowers users to create stronger passwords. This project not only promotes better password practices but also enhances overall online security by offering a proactive defense against weak and compromised credentials.

2.MILESTONES

The article titled “Multi-Class Classification Prediction Model for Password Strength Based on Deep Learning” by Seok Jun Kim and Byung Mun Lee proposes a deep learning-based model to evaluate password strength, addressing the limitations of existing evaluation methods. Traditional password strength indexes assess complexity based on length and character diversity but fail to account for leaked frequency, making them less reliable against breached passwords. To overcome this, the authors introduce a multi-class classification model that considers both password composition and leaked frequency. The model uses deep learning techniques to classify passwords into five categories based on their likelihood of being compromised. During the data preprocessing phase, the authors extract feature values such as ludsScore, zxcvbnScore, and levenshteinScore from password datasets, which represent complexity, pattern similarity, and edit distance, respectively. The model is trained and validated using labeled data, with leaked passwords categorized into different strength levels. The evaluation process confirms the model’s effectiveness, achieving 99.4% accuracy in identifying leaked passwords during testing. This solution enhances password strength evaluation by considering the frequency of password breaches, making it more reliable and practical for real-world applications. The authors conclude that integrating leak frequency into password strength analysis significantly improves security and offers a more robust defense against password-based attacks.

The article titled “Analyzing Password Strength: A Combinatorial Entropy Approach” by Naem Azam Chowdhury introduces a novel methodology for evaluating password strength by integrating combinatorial entropy calculations. The study highlights the limitations of traditional strength metrics, which primarily focus on password length and character diversity, by introducing combinatorial entropy as a more robust measure of password complexity. The proposed model systematically analyzes diverse password quality metrics, comparing their strengths and weaknesses. It employs clustering analysis to group passwords based on quality measures, revealing distinct categories such as weak, fair, medium, and strong. Additionally, the research presents experimental results demonstrating a positive correlation between password complexity and cracking difficulty. To enhance password resilience, the authors introduce an enhanced combinatorial entropy algorithm, which applies penalties for common patterns and evaluates the password's resistance against brute-force and dictionary attacks. The study also conducts rigorous experiments, applying statistical metrics such as Pearson correlation and Maximum Information Coefficient (MIC) to evaluate the effectiveness of the proposed model. The results confirm that combinatorial entropy offers a more accurate and reliable measure of password strength compared to traditional evaluation methods. This research provides valuable insights into developing more resilient authentication systems by incorporating advanced entropy-based approaches.

The article titled “Advancing User Classification Models: A Comparative Analysis of Machine Learning Approaches to Enhance Faculty Password Policies at the University of Buraimi” by Boumedyen Shannaq, Oualid Ali, Said Al Maqbali, and Afraa Al-Zeidi explores the effectiveness of various machine learning (ML) algorithms in strengthening password security. The study aims to enhance faculty password policies by classifying users based on password patterns, identifying weak and potentially compromised credentials. The authors evaluate five ML models: Neural Networks (NN), Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), using TF-IDF vectorization to transform passwords into numerical vectors. The models are assessed based on accuracy, precision, recall, and F1-score. The Neural Network model demonstrates the highest accuracy (0.129526) and F1-score (0.118785), indicating its superior ability to identify complex password patterns. The Random Forest model follows closely, offering strong performance in handling large datasets. In contrast, the K-Neighbors model ranks the lowest, struggling with precision and recall. The results confirm that Neural Networks and Random Forests are the most effective algorithms for password classification, offering valuable insights for building robust faculty authentication systems. The authors conclude that ML-based classification models significantly enhance password security by detecting patterns indicative of weak or reused passwords, thereby mitigating the risk of credential-based attacks.

The article titled “A Conceptual Framework for Assessing Password Quality” by Wanli Ma, John Campbell, Dat Tran, and Dale Kleeman presents a novel approach to measuring password strength using a Password Quality Indicator (PQI). The authors highlight that while password authentication is still the most widely used and cost-effective security mechanism, its effectiveness relies heavily on password strength. To address the limitations of traditional password guidelines, the authors propose the PQI framework, which evaluates password quality based on Levenshtein's edit distance and effective password length. The PQI metric quantifies how different a password is from dictionary words and calculates its equivalent strength in terms of cracking difficulty. The authors suggest that a good password should have a minimum Levenshtein edit distance of 3 from dictionary words and an effective password length of at least 14. To simplify this for users, they recommend a rule where strong passwords should be at least 8 characters long with at least 3 special characters plus other alphanumeric characters. This

simplified guideline offers an easy-to-remember yet effective rule for creating robust passwords. The study concludes that the PQI framework provides a practical, measurable, and effective method for assessing password strength, enhancing security practices, and guiding proactive password policies.

The article titled "Analysis of Default Passwords in Routers Against Brute-Force Attack" by Mohammed Farik and ABM Shawkat Ali examines the vulnerability of default router passwords to brute-force attacks. The authors highlight that router manufacturers provide default passwords for initial setup, but many users fail to change them, leaving networks susceptible to attacks. The study analyzes 1096 default passwords from various router models, evaluating their entropy, length, and cardinality to measure their resistance against brute-force techniques. Using PasswordStrengthCalculator.org, the authors determine that 96.4% of the default passwords can be cracked within 32 seconds, while only 0.0064% would require over 20 years to crack. The findings reveal that most default passwords have low entropy (below 52.4 bits), are shorter than the recommended 8-character length, and lack complexity, making them highly vulnerable. The authors recommend that router manufacturers implement built-in password strength enforcement mechanisms, requiring longer passwords with higher entropy and character diversity. They propose raising the minimum password standard to 12 characters with 94 cardinality to enhance security and withstand brute-force attacks. This study underscores the inadequacy of default router passwords and advocates for stricter password policies to prevent unauthorized access.

The article titled "A Probabilistic Framework for Improved Password Strength Metrics" by Javier Galbally, Iwen Coisel, and Ignacio Sanchez introduces a novel probabilistic model for evaluating password strength, addressing the limitations of traditional strength metrics. The authors propose two new statistical models: simple Markov chain (sMC) and layered Markov chain (lMC), which assign probabilities to passwords based on their likelihood of occurrence. These models leverage large public datasets of 75,000,000 real-world passwords to accurately assess password strength and resistance against guessing attacks. The sMC model evaluates the transition probability between consecutive password characters, while the lMC model considers the position of characters within the password, making it more accurate for strength assessment. The models were tested against password-cracking techniques, including brute-force and dictionary attacks, using John-the-Ripper. The results showed that 74% of passwords were cracked within 12 hours, confirming the vulnerability of common passwords. The models consistently assigned lower strength scores to easily cracked passwords and higher scores to stronger ones. The authors conclude that the probabilistic framework offers a more reliable and flexible method for evaluating password strength compared to traditional entropy-based metrics. The framework can be integrated into real-time password selection processes, helping users create stronger passwords and enhancing overall security against modern password-guessing attacks.

The article titled "Analyzing Password Strength" by Martin M.A. Devillers investigates the security of user-chosen passwords through an extensive empirical analysis of the RockYou! dataset, which contains 32 million plaintext passwords leaked from a social networking service. The study reveals alarming patterns, with over 90% of the passwords categorized as highly insecure. Through preliminary tests, the author examines the length distribution, character composition, and common patterns in the dataset. The results show that nearly 42% of passwords consist solely of lowercase letters, while 16% are purely numeric, making them highly vulnerable to guessing attacks. The author applies pattern analysis to identify common password structures, revealing that 15% of passwords consist of a word suffixed with the digit "1". Furthermore, an n-gram model is used to assess password predictability by calculating the occurrence probabilities of character sequences. The study concludes that trigrams (3-grams) offer the best trade-off between accuracy and performance for password strength estimation. The author develops a password checker that combines pattern analysis, dictionary checks, and n-gram scoring to assess password security. When tested on the dataset, 95% of the passwords were flagged as highly unsafe. The study highlights the poor password practices among users and underscores the need for stronger password policies, enhanced complexity requirements, and user education to mitigate security risks.

The article titled "Exploratory Data Analysis on Username-Password Dataset" by Vanita Jain, Rishab Bansal, and Mahima Swami investigates password security by performing Exploratory Data Analysis (EDA) on a 100 million email-password dataset. The authors analyze the password strength, character composition, and common patterns, providing valuable insights into user password practices. The study reveals that only 0.3% of the passwords are categorized as strong, while over 50% are weak, making them highly vulnerable to attacks. The analysis identifies seven major character set categories, with the most common being lowercase + numbers and lowercase + uppercase + numbers, accounting for 72% of the dataset. The most frequently used password is '123456', occurring 865,098 times (0.8% of the entire dataset), followed by other weak and easily guessable passwords like '1234567','000000' , and 'password1'. The study also identifies common domains associated with the email-password pairs, with 'yahoo.com' being the most frequent, representing 17% of the dataset. Additionally, the authors examine the occurrence of Unicode characters, revealing that the most common Unicode character 'a'

(ASCII value 1072) appears 1,942,387 times, accounting for 19% of all Unicode characters. The EDA further calculates the ratios of alphabetic letters, numeric digits, and symbols in the passwords, highlighting the overall simplicity of user-created credentials. The study concludes that the prevalence of weak and predictable passwords emphasizes the need for stronger password policies, improved security awareness, and the adoption of multi-factor authentication to mitigate the risk of credential-based attacks.

The article titled “Strength Analysis of Real-Life Passwords Using Markov Models” by Viktor Taneski, Marko Kompara, Marjan Heričko, and Boštjan Brumen explores the effectiveness of Markov models for evaluating password strength. The authors address the limitations of traditional password checkers by proposing a proactive password strength meter based on Markov models, which estimates the likelihood of a password being guessed. The study investigates whether one universal Markov model is sufficient for accurate password strength estimation or if multiple models trained on different datasets are necessary. The authors analyze 12 real-life password datasets, including RockYou, PhpBB, MySpace, and Faithwriters, using first-order Markov models trained on half of each dataset, while the remaining halves are used for testing. The results show that no single Markov model performs optimally across all datasets, confirming the need for multiple models to accurately assess the strength of diverse password groups. The study also reveals that the size and diversity of the training dataset significantly influence the model's accuracy. The RockYou and 10 Million Passwords datasets produce the most accurate models, highlighting the importance of large and varied datasets. The authors conclude that a multi-model Markov-based password checker provides more reliable strength evaluation, offering a practical solution for identifying weak passwords and enhancing password security.

The article titled “Deep Learning for Password Guessing and Password Strength Evaluation: A Survey” by Tao Zhang, Zelei Cheng, Yi Qin, Qiang Li, and Lin Shi provides a comprehensive overview of deep learning techniques used for password guessing and password strength evaluation. The authors highlight that traditional rule-based password guessing methods, such as Markov models and PCFG, are limited in handling complex password patterns. To overcome these limitations, the paper explores deep learning models, including Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Generative Adversarial Networks (GAN), and transfer learning approaches, which significantly improve the accuracy and efficiency of password guessing. The authors classify password guessing into two stages: password pattern extraction and candidate password generation, where deep learning models automatically learn complex patterns from large leaked password datasets. The survey also covers deep learning-based password strength meters, which outperform traditional methods by accurately evaluating password complexity and resistance to guessing attacks. The authors conclude that deep learning methods provide more accurate, practical, and scalable solutions for both password guessing and strength evaluation, making them superior to conventional approaches in real-world applications.

The article titled “Password Strength Verification Based on Machine Learning Algorithms and LSTM Recurrent Neural Networks” by Vladimir V. Belikov and Ivan A. Prokuronov presents a machine learning-based password strength verifier designed to improve the security of password authentication systems. The authors emphasize the vulnerability of traditional password verification methods to brute-force and dictionary attacks, highlighting the need for more advanced techniques. The proposed model uses supervised machine learning algorithms, including Support Vector Machines (SVM), Random Forest, Boosting, and a Long Short-Term Memory (LSTM) recurrent neural network. The TF-IDF method is applied during data preprocessing to convert passwords into vector representations, while cross-validation is used for hyperparameter tuning. The experimental results demonstrate that the LSTM model significantly outperforms traditional machine learning algorithms, achieving an accuracy of 99.94% and a macro F1-score of 99.92%. The authors conclude that LSTM networks offer a more reliable and scalable solution for password strength verification, making them particularly effective for enhancing password policies in various authentication systems.

The article titled “Machine-Learning-Based Password-Strength-Estimation Approach for Passwords of Lithuanian Context” by Ema Darbutaitė, Pavel Stefanovič, and Simona Ramanauskaitė presents a machine learning model designed to estimate the strength of Lithuanian-language passwords, addressing the lack of language-specific password-strength meters. The authors highlight that most existing models are tailored for English-language passwords, making them ineffective for less common languages. To bridge this gap, the authors compile a Lithuanian password dataset and integrate it with international password datasets to create a combined training set. The zxcvbn password-strength meter is modified by incorporating Lithuanian dictionaries and applying four similarity metrics: Fuzz, Levenshtein Jaro, Levenshtein Jaro Winkler, and Levenshtein Ratio. The modified model assigns passwords to five strength classes ranging from "too guessable" to "very unguessable". The authors evaluate the performance of five machine learning models: Naïve Bayes, k-Nearest Neighbors (kNN), Decision Tree, Linear Model, and Support Vector Machine (SVM). The Decision Tree model with Levenshtein Ratio similarity achieves the highest accuracy of 78%, demonstrating the model’s effectiveness in classifying

Lithuanian and international passwords. The results confirm that language-specific password-strength estimation is more accurate and reliable than traditional English-based meters, offering a scalable solution for other lesser-used languages.

The article titled "Measuring Password Guessability for an Entire University" by Michelle L. Mazurek, Saranga Komanduri, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Patrick Gage Kelley, Richard Shay, and Blase Ur presents an empirical study on password strength and guessability using 25,000 real-world passwords from Carnegie Mellon University (CMU). The study aims to evaluate the resistance of high-value passwords used for critical services, such as email, payroll, and course management, against offline guessing attacks. The authors employ statistical analysis and password cracking simulations to measure password strength, identifying correlations between demographic factors (e.g., gender, age, and college affiliation) and password robustness. The findings reveal that users from computer science and engineering departments create stronger passwords, while business school users tend to use weaker ones. The study also demonstrates that password complexity policies significantly influence password strength, with stricter policies resulting in 1.8 times stronger passwords. Moreover, the authors compare the CMU password dataset with leaked and simulated password sets, concluding that passwords from online studies more closely resemble real, high-value passwords than leaked, low-value account passwords. The study highlights the limitations of existing password strength meters and proposes using guessability metrics as a more accurate measure of password security. The authors conclude that demographic factors, password composition rules, and user sentiment significantly impact password strength, offering valuable insights for designing more effective password policies.

The article titled "Password Strength Analyzer Using Segmentation Algorithms" by Sivapriya K and Deepthi L.R introduces a segmentation-based approach for evaluating password strength by analyzing whether a password is based on user attributes, such as names, birthdates, or phone numbers. The authors highlight that traditional password strength meters focus only on length and character complexity but fail to detect patterns linked to personal information, making them less reliable. To overcome this, the authors propose three segmentation algorithms: Maximum Matching, Triangular Matrix, and the Password Segmentation Algorithm (PSA). The PSA is designed to optimally segment passwords into meaningful substrings and compare them with user attributes to detect correlations. Passwords with low correlation values to personal data are considered strong, while highly correlated passwords are marked as weak. The study demonstrates that PSA outperforms the other two algorithms in accurately detecting weak, user-related passwords by applying dynamic programming and optimal segmentation rules. The authors conclude that PSA-based password strength meters offer enhanced security by identifying and discouraging the use of easily guessable, user-specific passwords.

Dalton Gusaas, in his 2015 study titled *Password Strength Meters: Implementations and Effectiveness*, explores the functionality and reliability of various password strength meters (PSMs). This research examines different implementations, including rule-based meters, adaptive password strength meters (APSM), and analyzer modifiers for passwords (AMP). Rule-based meters, commonly used by organizations, rely on predefined complexity requirements but often fail to provide an accurate measure of password security. The APSM, utilizing n-gram models trained on leaked password datasets, offers a more data-driven approach to password evaluation. The AMP, employing probabilistic context-free grammars, not only assesses password strength but also suggests modifications to improve security. Gusaas concludes that while rule-based meters are straightforward to implement, they may inadvertently assist attackers by promoting predictable password structures. Comparatively, the APSM and AMP demonstrate superior performance by dynamically assessing and strengthening passwords based on real-world data. Among the three, the AMP proves to be the most effective in enhancing password security. This study highlights the need for future password strength meters to integrate adaptive learning and modification techniques for greater security resilience.

The article titled "Real-Time Password Strength Analysis on a Web Application Using Multiple Machine Learning Approaches" by Umar Farooq proposes a real-time password strength evaluation model for web applications, using multiple machine learning algorithms to classify passwords. The model integrates Decision Tree (DT), Naïve Bayes (NB), Linear Regression (LR), Random Forest (RF), and Neural Network (NN) classifiers to categorize passwords as Weak, Medium, or Strong. The Decision Tree model demonstrates the highest accuracy (99%), while Naïve Bayes achieves the lowest (87% accuracy). The model is tested using Burp Suite by simulating brute-force, dictionary, and reverse brute-force attacks on 250 accounts, where 86 out of 100 weak passwords are cracked, while none of the 150 strong passwords are compromised. The authors conclude that combining multiple machine learning models provides an effective solution for real-time password strength verification on web applications.

Margaret Nicole Todd, in her 2016 thesis titled *An Investigation of Machine Learning for Password Evaluation*, explores the potential of machine learning techniques in assessing password strength. Recognizing the limitations of traditional password strength meters, which often rely on predefined complexity rules, Todd investigates an alternative approach that evaluates passwords based on their structural uniqueness within a dataset. Using Kernel Density Estimation (KDE) and Principal Component Analysis (PCA) on a dataset of 10 million passwords, the study analyzes clustering patterns to differentiate weak passwords from strong ones. The research finds that traditional entropy-based methods do not fully capture password security and that machine learning models can more effectively identify structural similarities that make passwords predictable. While KDE shows promise in classifying passwords based on feature density, its scalability issues highlight the need for more efficient learning models. Todd's work suggests that an adaptive machine learning-based strength meter, which continuously learns from emerging password trends, could provide a more accurate assessment of password security. Future research should explore the integration of real-world password usage patterns and alternative machine learning models to refine password evaluation techniques.

Suganya G, Karpgavalli S, and Christina V published the article "Proactive Password Strength Analyzer Using Filters and Machine Learning Techniques" in the International Journal of Computer Applications (Volume 7, No. 14) on October 2010 [1]. This work proposes a framework to proactively analyze password strength by employing filters and Support Vector Machines (SVMs) to combat password vulnerabilities. This framework can be implemented as a submodule of the access control scheme. The password strength is tested using supervised machine learning algorithms, decision trees, and lexical analysis. Most of the survey results show that 10% to 15% of the user's passwords used mixed case, numbers, and symbols, but password is still the more compellingly authenticating the identity in many applications, hence the requirement of an effective password policy and proactive password checking system of an organization increases, that helps in selecting strong passwords and managing them, to protect the identity and the resources. Most of the password meters use lexical rules and the problem has been dealt using various filters and the support vector machine that outperformed other supervised machine learning algorithms namely oneR, C 4.5 Decision tree classifier, Multilayer perceptron and Naïve bayes classifier.

Mbaka, W.B. published the dissertation "An Online Neural Network Based Password Prediction, Generation, and Storage Scheme" at Strathmore University in September 2021 [1]. This dissertation sought to develop an online scheme to help Internet users generate stronger passphrases based on how predictable their preferred passwords are. Analysis of research findings asserted the need to incorporate neural networks, integrated data-driven insights, and derived concepts from the Markov chain model in the development of an online password predictive and generative scheme with an embedded password manager. The study analyzed existing literature on the character composition of human-created passwords, available tools for predictive analysis and generation of complex secret words and password managers. The resulting accuracy score after the scheme was trained using 50 epochs stood at 0.90332413, equivalent to 90.3%.

Gorle, B. published the article "Password Strength Analyzer" in the International Journal of Research Publication and Reviews, Volume 3, No. 11 in November 2022 [1]. Based on recent password security trends, this work presents a review of various algorithms and policies designed to foster strong passwords. The analysis relies on discussion of segmentation algorithms, analyses of password correlation to personal data, hashing and cryptographic strength measures. Users must be mindful of including alphanumeric characters, special symbols, etc. The strength of a password is checked through segmentation algorithms and analyses. Passwords with less correlation to personal details are the safest to use.

Sarkar, S. and Nandan, M. published the article "Password Strength Analysis and its Classification by Applying Machine Learning Based Techniques" in the 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA) [1]. This study presents the use of multiple supervised machine learning algorithms to classify password strength, highlighting the novelty of implementing XGBoost and Multilayer Perceptron in this context. Based on the experiment, the XGBoost outperformed other classifiers with an accuracy of 99%. The study focuses on a vital authentication mechanism and explains different service types using system passwords for authentication. Passwords are categorized as Weak, Medium, and Strong. Results indicated that machine learning approaches are sufficient to classify the passwords.

Aziz, E.F. and Baker, M.R. published the article "Enhancing Multi-Class Password Strength Prediction Through Machine Learning and Ensemble Techniques" in the International Journal of Safety and Security Engineering, Volume 14, No. 5, in October 2024 [1]. This research aims to explore a more advanced approach to password strength prediction that solves some of the existing shortcomings through a machine learning (ML) and ensemble model for multi-class classification. Here, in this research, we have employed Random Forest (RF), Decision Tree

(DT), Stochastic Gradient Descent (SGD), and Logistic Regression (LR) algorithms with Bagging and Stacking ensembling techniques. The results showed that the proposed approach provides a more accurate and versatile measure for password validation eradicating the problems encountered with the original method. The results proved the high efficiency of the used methods and showed more efficiency in prediction performance in comparison with the baseline models.

Arepalli, R., Supriya, G., Sheema, S., Devi, A., Govardhan, D., and Subbarao, G. published the article "Predicting The Strength of Password Using ML" in the International Journal of Innovative Research in Technology, Volume 11, Issue 2, in July 2024 [1]. This project aims to develop an accurate system capable of determining password strength based on linguistic properties, leveraging machine learning algorithms and Natural Language Processing in Python. This work will build a model to predict password strength. The outcome of this project can contribute significantly to the field of password security. By accurately predicting the password strength using NLP techniques we can assist users in creating stronger passwords and enhances overall cybersecurity.

Kuriakose, S., Teja, G.K., Duggi, S., Srivatsava, A.H., and Jonnalagadda, V. published the article “Machine Learning Based Password Strength Analysis” in the International Journal of Innovative Technology and Exploring Engineering (IJITEE) in July 2022 [1]. This study investigates the use of various machine learning methods such as Decision Tree (DT), Nave Bayes (NB), Logistic Regression (LR), and Random Forest (RF) on a web application in real time to force users to choose a secure password. The goal is to have the user's account being logged into if particularly the password strength from more than half of the algorithms is strong. [The study indicates the models' usefulness with algorithms but includes limited numerical data for comparison]. The proposed prototype implements numerous machine learning methods on a web application in real time to force users to choose a secure password.

Jared Wise, in his 2024 thesis titled *Enhancing Password Security and Memorability Using Machine Learning and Linguistic Patterns*, explores a novel approach to password generation that balances security and usability. Recognizing the challenges users face in creating strong yet memorable passwords, Wise proposes a machine learning-driven technique that utilizes linguistic patterns, particularly song lyrics, to generate secure yet easily recallable passwords. The study leverages large lyric datasets and natural language processing (NLP) to extract meaningful linguistic structures, employing transformer-based architectures to generate password phrases that maintain a balance between complexity and memorability. The research integrates recurrent neural networks (RNNs) to assess password security, measuring resilience against various attack strategies. Additionally, a user study evaluates the effectiveness of the generated passwords in terms of recall and security, demonstrating improvements over traditional password creation methods. By incorporating mnemonic aids such as narrative cues and personalized modifications, this approach enhances password retention while preserving robustness against attacks. Wise’s findings suggest that leveraging familiar linguistic structures can significantly improve password usability without compromising security. The study highlights the potential for future applications of artificial intelligence in cybersecurity and password management, advocating for user-friendly, AI-driven security solutions.

3.CONCLUSIONS

The Password Strength Analyzer using AI developed in this project offers a robust and efficient solution for evaluating and enhancing password security. By leveraging a hybrid LSTM-CNN model, the system accurately classifies passwords into Weak, Medium, and Strong categories based on their complexity and sequential patterns. The use of Conv1D (CNN) layers for local pattern extraction and LSTM layers for capturing long-term dependencies ensures reliable strength prediction. Additionally, the integration of the "Have I Been Pwned" API enables real-time leak detection, allowing users to verify if their passwords have been compromised in previous data breaches.The web-based application, built with HTML, CSS, and JavaScript, provides a user-friendly interface that offers instant feedback on password strength, score, and actionable recommendations. The backend, powered by Flask and TensorFlow, processes password evaluation requests efficiently and ensures seamless interaction between the frontend and the deep learning model.The system not only analyzes password strength but also promotes better password practices by offering customized recommendations, such as incorporating uppercase letters, digits, and special symbols. This encourages users to create stronger and more resilient passwords, thereby enhancing online security.Overall, the Password Strength Analyzer demonstrates the practical application of deep learning in cybersecurity, providing an effective tool for password evaluation, real-time breach detection, and security enhancement, making it a valuable solution for individuals and organizations aiming to strengthen their authentication mechanisms.

4.REFERENCES

[1]. Multi-Class Classification Prediction Model for Password Strength Based on Deep Learning -Seok Jun Kim1, Byung Mun Lee1 - march 07 2023
[2]. Analyzing Password Strength: A Combinatorial Entropy Approach - Naem Azam Chowdhury1 - 11 January 2024
[3]. Advancing user classification models: A comparative analysis of machine learning approaches to enhance faculty password policies at the University of Buraimi - Boumedyen Shannaq1, Oualid Ali2, Said Al Maqbali1, Afraa Al-Zeidi1 - 10 October 2024
[4]. A Conceptual Framework for Assessing Password Quality - Wanli Ma, John Campbell, Dat Tran, and Dale Kleeman - January 5 2007
[5]. Analysis Of Default Passwords In Routers Against Brute-Force Attack - Mohammed Farik, ABM Shawkat Ali- September 2015
[6]. A probabilistic Framework for Improved Password Strength Metrics - Javier Galbally, Iwen Coisel, Ignacio Sanchez - 30 march 2016
[7]. Analyzing Password Strength - Martin M.A. Devillers - July 2010
[8]. Exploratory Data Analysis on Username-Password Dataset - Vanita Jain1, Rishab Bansal2 and Mahima Swami - DOI: 10.5281/zenodo.5169881 - may 10 2021
[9]. Strength Analysis of Real-Life Passwords Using Markov Models - Viktor Taneski , Marko Kompara, Marjan Hericko and Boštjan Brumen - 30 september 2021
[10]. Deep Learning for Password Guessing and Password Strength Evaluation, A Survey - Tao Zhang134, Zelei Cheng2, Yi Qin, Qiang Li, Lin Shi - December 2020
[11]. Password strength verification based on machine learning algorithms and LSTM recurrent neural networks - Vladimir V. Belikov,
Ivan A. Prokuronov - 2 may 2023
[12]. Machine-Learning-Based Password-Strength-Estimation Approach for Passwords of Lithuanian Context - Ema Darbutaite, Pavel Stefanovic and Simona Ramanauskaite - 30 june 2023
[13]. Measuring Password Guessability for an Entire University - Michelle L. Mazurek, Saranga Komanduri, Timothy Vidas, Lujo Bauer,
Nicolas Christin, Lorrie Faith Cranor, Patrick Gage Kelley , Richard Shay, and Blase Ur - 13 november 2013
[14]. Password Strength Analyzer Using Segmentation Algorithms - Sivapriya K , Deepthi L.R -November 23 2020
[15]. Password Strength Meters: Implementations and Effectiveness - Dalton Gusaas - December 5 2015
[16]. Real Time Password Strength Analysis on a Web Application Using Multiple Machine Learning Approaches - Umar Farooq - December 2020
[17]. An Investigation of Machine Learning for Password Evaluation - Margaret Nicole Todd - November 2016
[18]. Proactive Password Strength Analyzer Using Filters and Machine Learning Techniques - Suganya G , Karpgavalli S, Christina V - October 2010
[19]. An Online neural network based password prediction, generation and storage scheme. - Mbaka, Winnie Bahati - September 2021
[20]. Password Strength Analyzer - Bhavani Gorle - November 2022
[21]. Password Strength Analysis and its Classification by Applying Machine Learning Based Techniques - Sakya Sarkar , Mauparna Nandan - august 2022
[22]. Enhancing Multi-Class Password Strength Prediction Through Machine Learning and Ensemble Techniques - Enas F. Aziz , Mohammed Rashad Baker - 15 october 2024
[23]. Predicting The Strength of Password Using ML - RAJESH AREPALLI, G. SUPRIYA, SK. SHEEMA, A.L.S. DEVI, D. GOVARDHAN, G.
SUBBARAO - July 2024
[24]. Machine Learning Based Password Strength Analysis - Sony Kuriakose, G Krishna Teja, Sravan Duggi, A Harshel Srivatsava, Venkat Jonnalagadda - July 2022
[25]. Enhancing Password Security and Memorability Using Machine Enhancing Password Security and Memorability Using Machine Learning and Linguistic Patterns Learning and Linguistic Patterns - jared wise - december 2024