MTU

Ollscoil Teicneolaíochta na Mumhan

Munster Technological University

Computer Science Final Year Project

2023/24 Poster Selection

Nexus Centre,

MTU,

Bishopstown, Cork

May 2024

High-Tech Augmentative and Alternative Communication App for Children with Autism

Alison Adriaansen
BSc Honours Software Development
Department of Computer Science
MTU Cork, May 2024

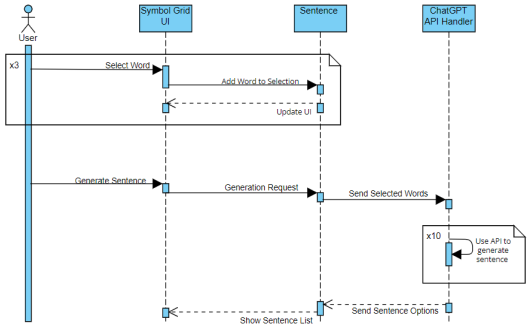
Abstract

Augmentative and Alternative Communication (AAC) offers a diverse array of communication methods for children affected by Autism Spectrum Disorder who experience limited speech or are nonverbal. Instead of relying on conventional spoken language, AAC serves as a valuable means through which these children can express themselves and engage in meaningful communication. Can artificial intelligence be leveraged to improve the efficiency and speed at which users of AAC can achieve their desired expression? This question is tackled by this project through academic research and development of a proof-of-concept high-tech AAC mobile application.

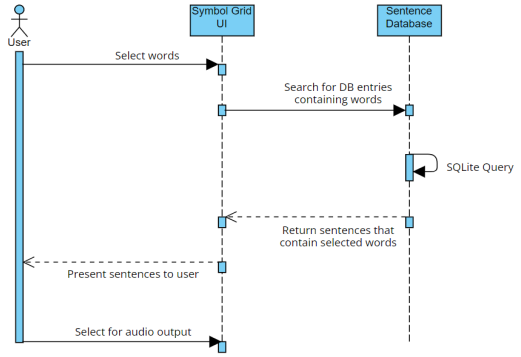
Contributions

- Low-cost Android solution for the AAC app market
- Showcases the benefits of Large Language Model integration in AAC
- Research-backed user interface design for enhanced AAC usability

Use Cases



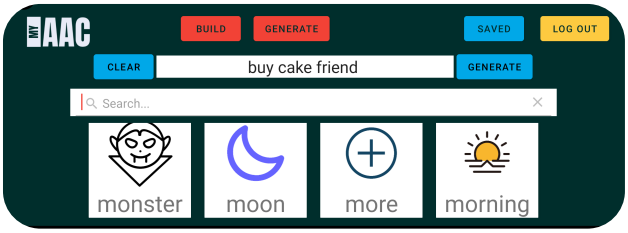
1. Leverage LLM for sentence prediction



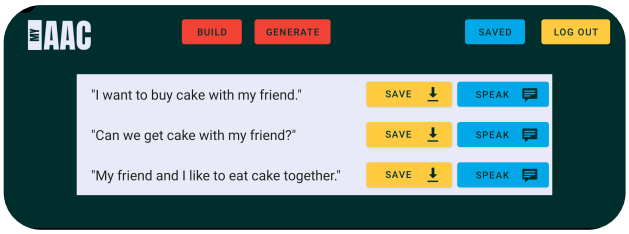
2. Algorithm for sentence database search

Application

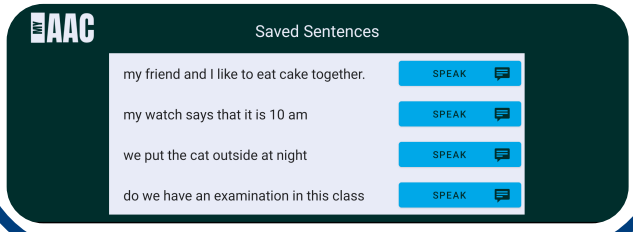
Symbol Grid UI



Sentence Prediction Options



User-Specific Saved Sentences



Tools & Technologies

Android Studio, Android, Java, SQLite, Firebase Auth, Firebase Firestore, OpenAI

Acknowledgments

The author would like to acknowledge the Department of Computer Science with special thanks to Dr. Ignacio Castiñeiras for supervising this project.

Machine Learning for Malicious URL Identification and Browser Integration

Eman Al Badri, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Problem Statement & Objectives

How can machine learning-based malicious URL detection be integrated into web browsers to enhance user protection?
Phishing remains to be one of the top social engineering attack methods criminals use to gain initial access, highlighting that 82% of cyberattacks are initiated through phishing [1]. This project explores URL analysis, forms of malicious URL detection, and the effectiveness of anti-phishing training. The proposed solution is to integrate a Machine learning model to a simple Chrome Extension designed to block a malicious webpage and educate the user on phishing.

Implementation Flowcharts

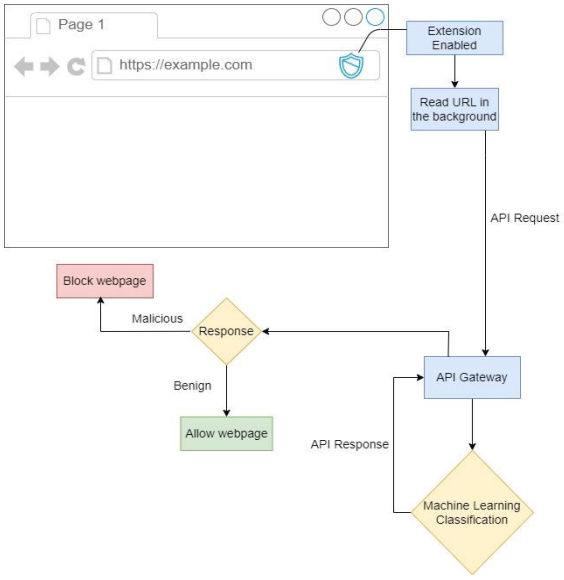


Figure 1: Flowchart describing the functionality of the chrome extension

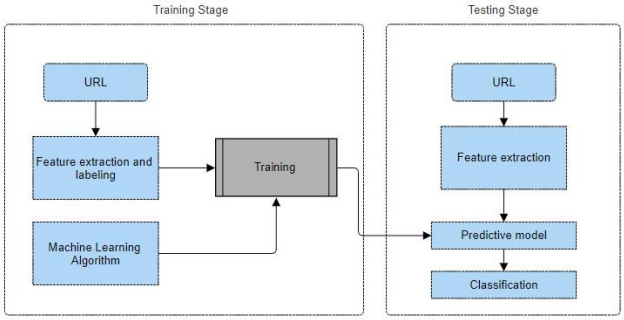


Figure 2: Flowchart describing the functionality of the Machine Learning Algorithm

Deep Learning vs Machine Learning Evaluation

14 Lexical features were extracted from the dataset for the models, these include:

- Number of Dots
- Number of Dashes
- Number of Underscores
- Number of Percent
- Number of Hashes
- Number of Ampersands
- Tilde Symbol
- Number of Numbers
- URL Length
- At Symbol
- IP Address
- Hostname Length
- Query Length
- Https

Classification	Precision	Recall	F1-score
Benign	97%	100%	98%
Malicious	100%	95%	97%
Accuracy			**98%**

Table 1: Classification Report for **Keras Sequential Model (KSM)**

Classification	Precision	Recall	F1-score
Benign	98%	99%	98%
Malicious	99%	97%	98%
Accuracy			**99%**

Table 2: Classification Report for **Random Forest Classifier (RFC)**

- The difference between both models in terms of metrics is very small. For that reason, prediction speed was considered & analysed:
  - Prediction Time for RFC: 0.078 s
  - Prediction Time for KSM: 5.038 s
- Overall, the Random Forest Classifier has shown to provide better results in terms of time and prediction. Therefore, it was more efficient for application.

Conclusion & Future Work

This project focuses on lexical-based features of a URL, neglecting behavioural aspects of URLs and webpage content. Furthermore, security concerns regarding web extensions are noted, as they may introduce vulnerabilities despite this extension's defensive potential. Future considerations include implementing security measures for web extensions & implementing content-based features for enhanced accuracy.

Technology Used

References

1. ENISA, "Enisa threat landscape 2023," https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023, ENISA, Tech. Rep., 2023, accessed: 30/10/2023

Acknowledgments

I would like to thank my supervisors Dr Alison O'Shea and Dr Kashif Ahmed for their guidance on the project.

First Person Puzzle Game

With database integration

Daniel Borg, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

This project was born out of my passion for Video Games as a medium, along with the desire to attempt to merging surveying of data from player interaction with digital software. I believe that interactive entertainment can be a very powerful tool in the field of research into people's reactions and solutions to certain problems, with this project aiming to replicate that effect.

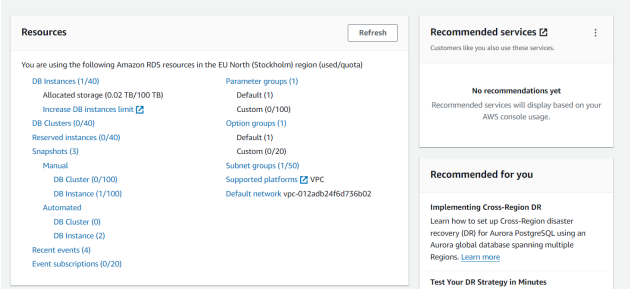
Specifications

This project has been made utilizing two major parts, an Online AWS hosted MySQL database through which the data is stored and kept, and the game itself, developed in Unreal Engine 5 and communicating with the database through integrated plugins that allow direct data transmission between Unreal Engine and a MySQL database system.

The Database and Analytics

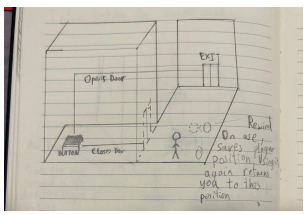
MySQL is a database structure commonly used throughout various systems, based in SQL syntax. The database used for this system utilizes two primary tables, one for each level and one for each possible solution. The solutions include information in regards to resources and abilities used by the player to beat the level.

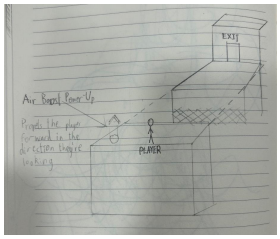
In terms of what details are stored, the players time taken is tallied, as well as the amount of times they used each of the game's four powerups individually. These details are then tallied through an HTTP request to the database. Currently, the REST service will be hosted locally as a proof of concept, though the MySQL database itself is hosted using AWS's free trial.



The Game

The game is developed in Unreal Engine 5, set from the first person perspective and inspired by other first person puzzle games, such as The Talos Principle and Portal. The player uses movement mechanics to navigate through the levels and reach their end, at which point the statistics of the player's solution will be added to the database as a valid completion point. The game was also evaluated and designed with the works of Jesse Schell's Lenses in mind, following the practices outlined for designing puzzles as outlined in his book "The Art of Game Design"





Conclusions

- Whilst the end result of the project as shown in this presentation is rudimentary, it acts as a proof of both concept and function, the ability to make this sort of system in a game and other functions that could be possible. The game's solution storing system can work at a higher level for potential psychological studies, or at a much smaller scale as a social game aspect, with players trying to get better scores than one another.

References

1. Jesse Schell, "The Art of Game Design"
2. Gallup, Jennifer, and Barbara Serianni, " Developing Friendships and an Awareness of Emotions Using Video Games: Perceptions of Four Young Adults with Autism."

Roster Optimisation for at Home Care

Josh Cronin, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

The goal of this project is to develop a tool that can be utilised to generate rosters for nurses and carers providing at home care while allowing the carers to provide their own preferences to the roster itself to ensure that they are also happy with it.

Motivation for roster optimization for at home care

The problem of roster generation is a classical optimization task, which can be tackled in a number of ways by machine learning. The task of generating an optimal roster is a complex issue which is very time consuming to manual create, which is why more and more tools are being developed in order to facilitate making them.
At home care rostering shares many similarities to the nurse rostering problem which has been tackled with a variety of approaches[1, 2, 4, 5], which work to varying levels of success, but all out preform the manual alternative in both time and satisfaction. The two approaches that were determined to show the most promise for the task were genetic algorithms[1] and reward machines[3]

Genetic Algorithm VS Reward Machine

Reinforcement learning, a subset of machine learning, offers a promising approach to optimization tasks. With its foundation in animal psychology, reinforcement learning involves an agent interacting with an environment, receiving rewards based on its actions.
Genetic algorithms, another popular optimization technique, represent solutions as genomes and evolve optimal solutions through crossover and mutation operations. They leverage populations of solutions and fitness functions for evaluation, mimicking biological evolution.

"One Max" Problem

The "One Max" problem involves finding a binary string of maximum length where all bits are set to 1.

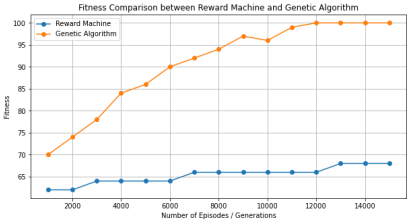


Fig 1

Figure 1 shows the comparative performance of a genetic algorithm and a reward machine for the "one max" problem. They are compared on score vs number of generations.

Roster Optimization

Constraint	Score
Missing Shift	0
Over Assigned Monthly	0
Illegal shift	0
Preferred Pattern	+10/-5
Preffered Shift	+50/-10
Under Assigned Average(Monthly)	-100
Fairly Assigned Average(Monthly)	+10
Under/Over Assigned Average(Weekly)	-10
Fairly Assigned Average(Monthly)	+5

Table 1 – Constraints and scores

Table 1 shows the various constraints currently being considered by the reward function and the score assigned to them. The higher the score the higher value meeting the constraint is worth. Hard constraints offer a higher reward to ensure they are met first.

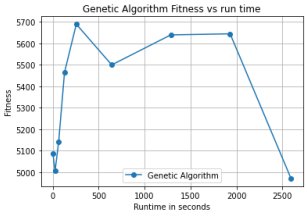


Fig 2

Hard constraints must be adhered to as they are based in labour law.
Soft constraints are maximised and are the preferences of carers

Figure 2 shows the performance of the genetic algorithm against the time it took to generate the result.
The genetic algorithm relies on mutation to locate unexplored solution spaces which can lead to lower fitness scores after longer runs or more generations as seen in the end of the graph

Future Work

- Fine tuning of the values of the reward function so that the resultant roster is optimal
- Ability to easily add and remove soft constraints on demand
- Sending personal rosters to each individual carer instead of requiring them to check an overall roster
- Finding a way to combine the consistency of a reward machine with the results of a genetic algorithm.

References

1. Muhammad, N, Shuaibu, A, Shuaibu, A, Shu'aibu, A, Idris, Y, Lawal, M, "Solving Nurse Rostering Problem Using Genetic Algorithm", *2021*.
2. Kyngäs, N, Nurmi, K, Ásgeirsson, E, Kyngäs, J. "Using the PEAST Algorithm to Roster Nurses in an Intensive-Care Unit in a Finnish Hospital"
3. Toro Icarte, R, Valenzano, R, Mcilraith, S "Reward Machines: Exploiting Reward Function Structure in Reinforcement Learning", 2022.
4. Hadwan, M, "Annealing Harmony Search Algorithm to Solve the Nurse Rostering Problem", 2022.
5. Leung, F, Lau, Y, Law, M, Djeng, S. "Artificial intelligence and end user tools to develop a nurse duty roster scheduling system" , 2022

Acknowledgments

The author would like to acknowledge the support of the MTU Department of Computing.

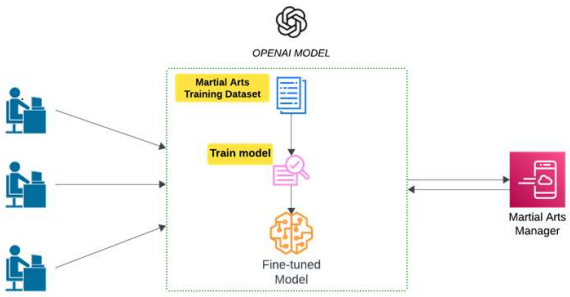
MAM
Martial Arts Manager with LLM-powered backend fine-tuned using a custom-Taekwondo corpus

Christopher Duffy
BSc Honours in Software Development
Department of Computer Science
MTU Cork, May 2024

Introduction

The objective of this project is to develop a mobile application using Java, XML and a Firebase Realtime Database. The aim of this application is to allow users to have a fast and secure method to transfer important information, to view useful data like the timetable and interact with a chatbot for useful information. The project looks to tackle a problem relating to the use of a lot of paper instead of having a method that can do this online for a local club.

System Architecture



Project Goals

The goal of the project is to create a Graphical User Interface (GUI) to allow a user to engage with a variety of features [1].

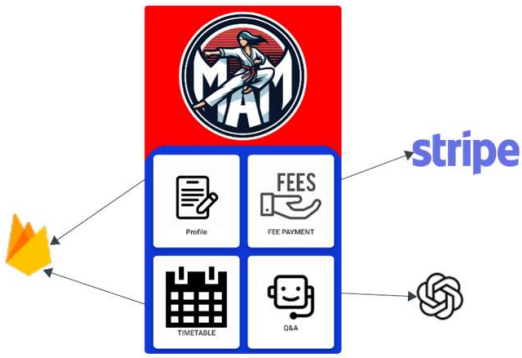
Using modern technology to register, make fee payments, see the timetable and use the chatbot allows for faster and more secure storage of important information that can be stored or viewed.

The use of a fine-tuned custom LLM for the chatbot enables club members to ask questions relating to training and diet which are specific to martial arts.

Features

1. **Profile**
   - Update profile information
2. **Fee Payment**
   - Card Payment System
   - Stripe Payment System
3. **Timetable**
   - View class timetable
4. **AI Assistant**
   - Custom AI assistant chatbot
   - Fine-tuned using OpenAI
5. **Multiple User Roles**
   - Admin
   - Instructor
   - Student

Operations of MAM



Technologies Used

References

1. Stuurman, Sylvia and van Gastel, Bernard E and Passier, Harrie JM, The design of mobile apps: what and how to teach?, Proceedings of the Computer Science Education Research Conference, 2014.

Acknowledgments

The author would like to thank Dr Larkin Cunningham for support on the report and Dr Seamus Lankford for supervising the practical work.

The Use of Smart Device in a Smart Home

Jordan Flynn, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024

Introduction

The modern world can be a struggle on the best of days for most people, however for people with disabilities this struggle can be unbearable, whether those disability are physical or mental and for some it is both. The aim of my project is to help these people continue to live independently using a web of devices all connected with a shared software for easy integration and scaling. Internet of Thing devices also known as smart devices allow for this aim to be achieved.

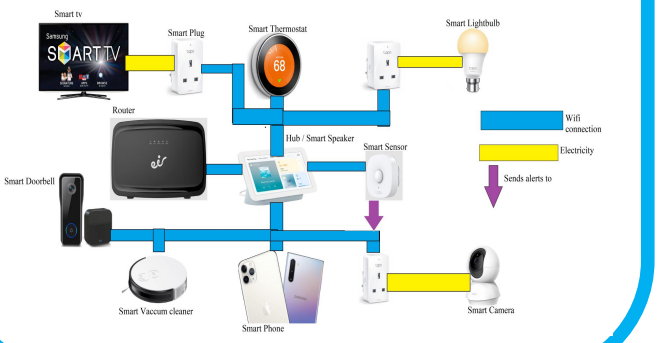
Goals of the Project

The main goal of my project to help my grand aunt. I will achieve this by automating her daily tasks with the use of smart devices and creating a "touchless home" with an aim to achieve an automation amount of 70% automation.

Also making a website to contain a brief overview of the project while also having links to the products so that they may purchased

List of Devices Used

Google Nest
XTU Smart doorbell
Tapo Smart bulbs
Tapo Smart plugs
Tapo Smart motion sensor
Tapo Vacuum cleaner
Samsung Smart Phone
Samsung Smart TV
Nest Smart Thermostat



QR Code to my Website

SCAN ME

Conclusions

- While Smart Devices can't fully automate everything, the tasks they do automate allows for easier lives for everyone.

Acknowledgments

I would like to acknowledge, my mother for proof reading and all her hard work and my grand aunt with her reviews of the devices.

Brain Tumour Detection using Image Classification

Oisín Flynn, BSc Honours in <Software Development>
Department of Computer Science,
MTU Cork, May 2024

Introduction

Early detection of brain tumours is crucial in improving the prognosis and treatment outcomes for patients. Current diagnostic methods, like Magnetic Resonance Imaging (MRI) scans, require expert analysis and are often time-consuming, with radiologists' principal diagnosis being 69% [1]. By leveraging advanced image classification techniques, using Convolutional Neural Networks (CNNs) this project aims to automate and enhance the accuracy of brain tumour detection, potentially reducing diagnostic times and human error, significantly impacting patient care.

Comparing CNNs & Visual Cortex

V1: Interprets basic visual details like light and shadow.
V2: Organises these details into shapes.
V3: Tracks movement.
V4: Identifies colours and object forms. [2]

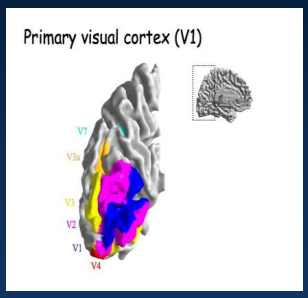
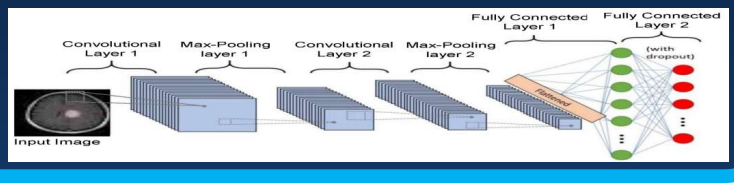


Fig 1: Map of Human Visual Cortex [2]

Convolutional Layers: Like V1, these layers identify basic visual cues edges and textures. **Pooling Layers:** Simplify and shrink the input from convolutional layers, much like how V2 starts shaping objects. **Fully Connected Layers:** Integrate all the information to classify images, like how V3 and V4 discern movement and complex forms, including colour.

Fig 2: Architecture of CNN [3]



Dataset & Classes

The dataset from Kaggle consisted of 7023 images of human brain MRI scans classified into 4 classes.

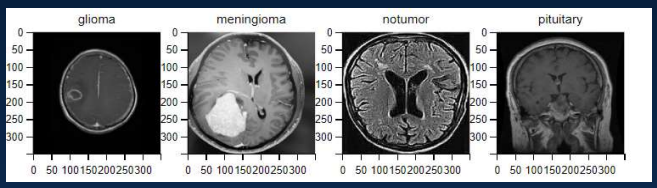


Fig 3: Example of Each Class of Brain MRI Image

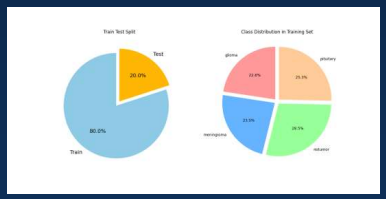


Fig 4: Train Test Split & Class Distribution of Training Images

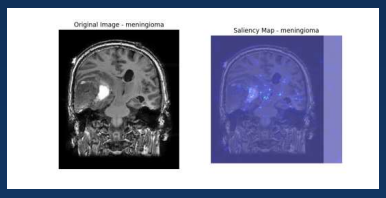


Fig 5: Saliency Map of Meningioma Prediction

EfficientNetB1 Predictions

I used EfficientNetB1 model with transfer learning which achieved 98% accuracy.

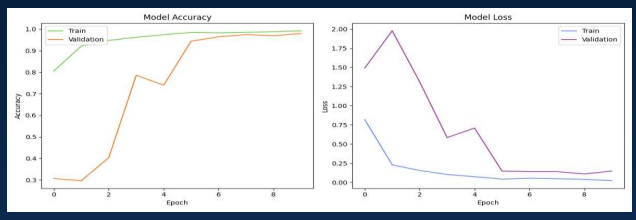


Fig 5: EfficientNetB1 Model Accuracy & Loss

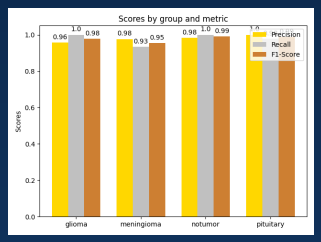


Fig 6: Metrics For Each Class

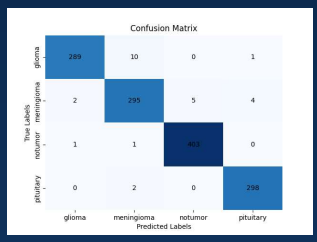


Fig 7: Confusion Matrix

The findings demonstrate that CNNs can significantly enhance the accuracy of brain tumour diagnosis with a 29% increase in accuracy from radiologists' principal diagnosis.

Conclusions

Future developments could include expanding the dataset to include a broader range of tumour types and stages, refining the CNN architecture to improve accuracy and reduce false positives/negatives, and integrating the model into a clinical setting for real-world validation and feedback.

References

1. K. A. Manias et al., "Diagnostic accuracy and added value of qualitative radiological review of 1H-magnetic resonance spectroscopy in evaluation of childhood brain tumors," Neuro-oncology practice, vol. 6, no. 6, pp. 428–437, May 2019, doi: https://doi.org/10.1093/nop/npz010.
2. "Perception Lecture Notes: LGN and V1," Nyu.edu, 2024. https://www.cns.nyu.edu/~david/courses/perception/lecturenotes/V1/lgn-V1.html
3. M. Aloraini, A. Khan, Suliman Aladhadh, S. Habib, M. F. Alsharekh, and M. Islam, "Combining the Transformer and Convolution for Effective Brain Tumor Classification Using MRI Images," Applied sciences, vol. 13, no. 6, pp. 3680–3680, Mar. 2023.

Analysis of Guloader's Evasion Tactics

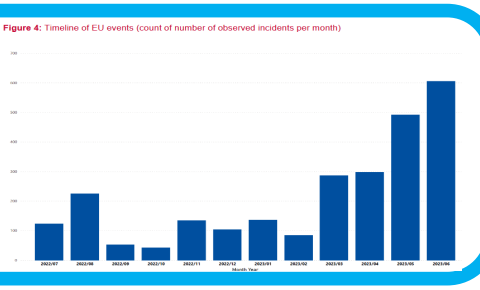
Shatrunjaya Manohar Gemawat
BSc Honours in IT Management
Department of Computer Science
MTU Cork, May 2024

Introduction

•Malware is one of the primary cybersecurity threats identified between 2022 and 2023.
•One of the biggest malware threats is still information stealers; e.g. Agent Tesla, Redline Stealer and FormoBook.
•Modern Malware has started to utilize anti-analysis techniques to evade detection and analysis.

Anti-Analysis, AKA Evasion Techniques

- Evasion techniques are becoming more common and creative in modern malware.
- This project aims to investigate these anti-analysis components within Guloader and provide countermeasures for these components to aid in Guloader's malware analysis.



Guloader

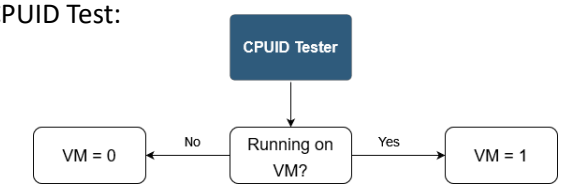
- Guloader is a loader-type malware that infiltrates devices to deliver malicious payloads.
- Various anti-analysis techniques have been observed in 6 different Guloader samples, such as anti-VM, anti-Debugging, anti-antivirus, anti-sandboxing, obfuscation, etc.

Implementation

- 6 different samples were taken of Guloader, in which 3 different payloads were observed, each utilizing different anti-analysis techniques.
- Joe's Sandbox, an automated sandboxing tool was used to analyze the intricate actions performed by these samples.
- The project concluded with the compilation of a list of anti-analysis techniques observed, together with remediation methods for each.
- Two supporting programs were developed to assist in implementing these countermeasures.

Countermeasures for the Evasive techniques

CPUID Test:



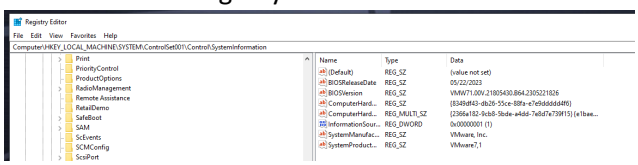

```
D:\MTU\Year 4\Final Year Project>cpuid_test
EAX is 0x40000000
Is VM: 0
```

```
C:\Users\shatr\OneDrive\Desktop>cpuid_test
EAX is 0x40000000
Is VM: 1
```

PowerShell Script:

```
PS C:\Users\shatr\OneDrive\Desktop> .\Registry_VM.ps1
Enter the term to search for in the registry: vmware
HKEY_LOCAL_MACHINE\SOFTWARE\Classes\Applications\VMwareHostOpen.exe
HKEY_LOCAL_MACHINE\SOFTWARE\Classes\Applications\VMwareHostOpen.exe\shell
HKEY_LOCAL_MACHINE\SOFTWARE\Classes\Applications\VMwareHostOpen.exe\shell\open
HKEY_LOCAL_MACHINE\SOFTWARE\Classes\Applications\VMwareHostOpen.exe\shell\open\command
```

VM Artifacts in Registry:



Conclusions

Guloader typically employs a range of evasion techniques (e.g. anti-VM, anti-sandboxing, etc.) to thwart analysis. This project investigated a range of Guloader samples and developed a set of guidelines for circumventing their evasion techniques, allowing successful analysis of the malware.

References

1. I. Lella et al., "ENISA THREAT LANDSCAPE 2023 ABOUT ENISA EDITORS," 2023. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023/@@download/fullReport

Acknowledgments

I would like to acknowledge my supervisors Hamdan Awan and David Murphy for their assistance with this project.

Developing a Horror-Themed Minecraft Mod via In-Game Assets and Mechanics

Alberto Gonzalez, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

The goal of this project was to develop a modification for the videogame Minecraft (a survival/construction sandbox game), with a strong focus on horror and survival on a brand new and hostile environment.

New World Generation and Structures

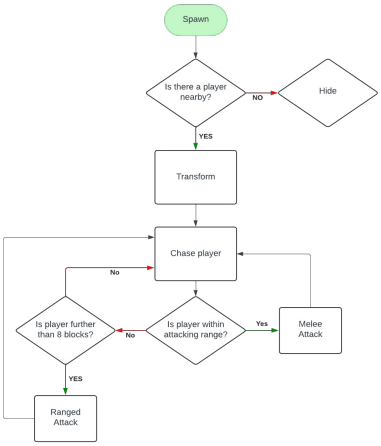
The mod adds a new procedurally generated dimension that serves as the backdrop for the array of newly added enemies. This dimension is a "corrupted" mirror of the normal world, with a very similar layout and new additions such as crystalline formations which yield new materials.

In the normal world, players can also find new structures that serve as gateways to this new dimension.

New AI-Driven Entities

A variety of new zombie-themed enemies have been added to the game, encompassing explosive variants, airborne counterparts, and those with the ability to transform themselves into objects to trick players.

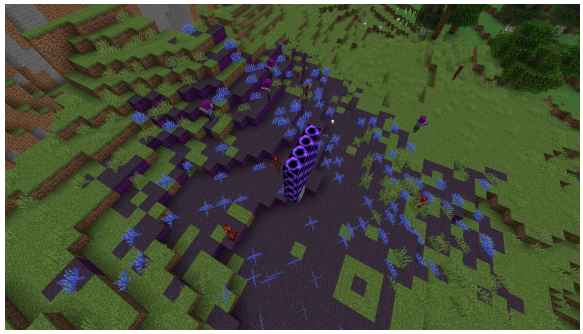
These enemies use some of Minecraft's own AI systems as well as brand-new ones built from the ground up specifically for these new additions.



New Mechanics

Portals may occasionally appear throughout the normal world that will gradually infect the world, slowly spreading and transforming the surrounding area into an undead wasteland.

The more area is infected, the more monsters appear and the more difficult it is to shut down the portal, so players must be swift in dealing with these anomalies or risk having their world infected and overrun by monsters.



User Reception and Future Work

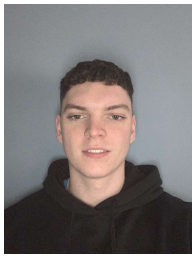
Engagement during play testing of the Alpha version of the mod was positive, with users providing valuable feedback to achieve a balanced and engaging experience. In the future, the mod will be uploaded online so that everyone may be able to experience all this brand-new content.

Acknowledgments

The author would like to acknowledge Christian O'Leary for supervising this project, alongside all the testers that played the Alpha version and gave their feedback to help make the best version of the mod.

DDoS Mitigation Using Load Balancing

Luke Hanley, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024



Introduction

Distributed Denial of Service (DDoS) attacks can harm online businesses, interfere with essential services, and disrupt user experience. The primary goal of this project is to research a strategy that allows service providers to keep their web pages or applications up and running even when a DDoS attack is in progress.

DDoS Attacks & Mitigation

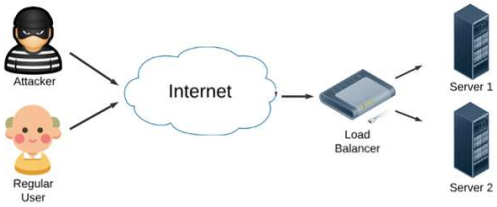
DDoS attacks can be classified into three different sections [1,2]: Flood attacks, Application layer attacks and Protocol based attacks. Each having their own distinct methods of attack.

Mitigation methods exist however, currently there is no sure-proof way to stop these attacks. Methods include rate-limiting, source IP blocking, as well as load balancing. In this project I have focused on load balancing as it is effective at eliminating a single-point of failure by distributing traffic across multiple servers.

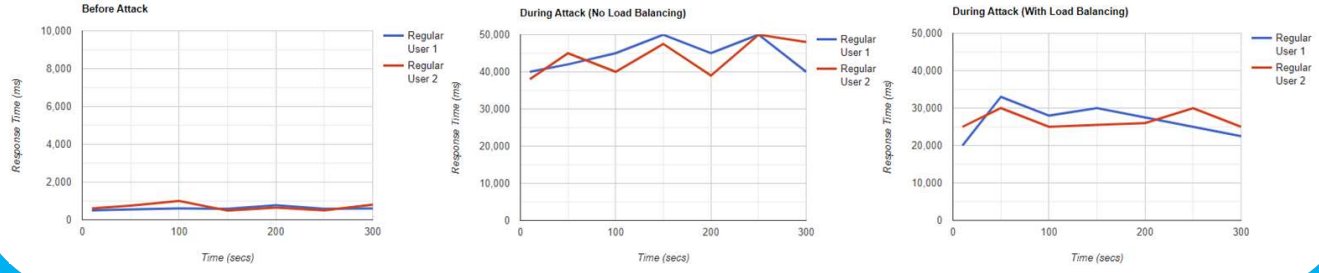
Methodology

In my project, I simulated a DDoS attack using GNS3. In the simulation I had:

- Three attackers to simulate the use of a botnet.
- Two regular users to test response times before and during the attack.
- Multiple different load balancing methods.
- Detection system looking for high numbers of requests from single IP addresses.



From the tests we can see that load balancing is a useful DDoS mitigation technique however still has its flaws. As seen in the diagrams below, response times improved slightly but the attack could not be stopped completely. This simulation proved that load balancing can help but must be used in conjunction with other mitigation methods to ensure a complete DDoS mitigation plan.



Conclusions

Load balancing proves be an effective DDoS mitigation technique. Although it cannot stop a DDoS attack on its own, when used as part of a layered security approach it can get rid of the single point of failure and help find a permanent solution for the DDoS attack.

References

1. Gupta, Brij B and Dahiya, Amrita, “Distributed Denial of Service (DDoS) Attacks: Classification, Attacks, Challenges and Countermeasures”.
2. Douligeris, Christos and Mitrokotsa, Aikaterini, “DDoS attacks and defence mechanisms: classification and state-of-the-art”.

Acknowledgments

Roland Katona
Noreen Gubbins
Dylan Smyth

Implementing an OT Test Environment for a Leading Pharmaceutical Company.

Fionn Healy, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024

Introduction

- Pharmaceutical manufacturing relies on highly available Operational Technology (OT) to provide safe, secure and reliable manufacturing systems.
- Safe testing and evaluation of changes to these systems requires dedicated test environments.

Problem statement

- The lack of an OT test environment in Pfizer Ringaskiddy results in the build of temporary servers for testing purposes, resulting in:

Inefficient Use of Time

Time consuming build of servers

Unreliable & Inaccurate Testing

- The reliability and availability of current development systems is put at risk as a result.

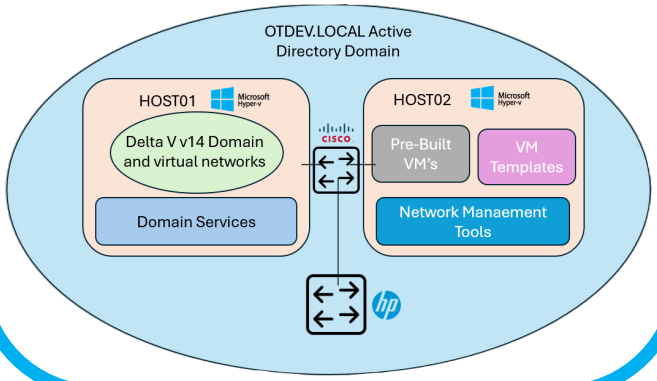
Proposed Solution Overview

- A dedicated test environment for use by the Automation Infrastructure Team is to be designed and deployed.
- Functional Requirements:
  - Virtual Machines running production Operating Systems.
  - Virtualised replica of a current Delta V "Dev" Distributed Control system.
  - Installation of the system into a datacentre environment.
  - Segmented from the surrounding datacentre systems.

Implemented Solution

A test environment was built and deployed with the following components:

- Delta V v14 DCS Environment
- Microsoft Hyper-V Environment.
- Physical and virtual networking.



Conclusions

- An OT test environment was delivered successfully to the Automation Infrastructure Team.
- Test and development activities have become more efficient and more reliable
- 3 of the 4 functional requirements were met successfully.
- Datacentre installation was not possible due to time constraints.

Acknowledgements

The author would like to acknowledge the following people:
- Project supervisor Mr. Roland Katona.
- Automation Infrastructure Team, Pfizer Pharmaceuticals, Ringaskiddy.

Geo-insight Microservice Orchestration

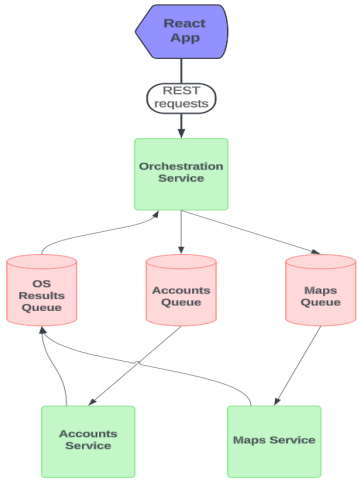
Cian Herlihy
BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Background and Motivation

The surge of data from online platforms presents a challenge in extracting and handling of huge data flow and demand. This project leverages microservices architecture(MSA) to move data between services fast and efficiently. With the integration of Google Maps API. MSA's capability is used to handle large scale data while providing geographical insights.

Geo-Insights

Overview



- Consumers interact through React front-end interface by submitting forms and viewing data
- Orchestration services redirects messages to correct service and acts as 1 stop for external communication
- Services communicate with Orchestration service using RabbitMQ queues for fast data transfer.

Technologies Used

Conclusions & Future Improvements

- Kubernetes Scaling with replication of docker containers works seamlessly for load balancing
- RabbitMQ offers fast messaging with a bonus of multi-threading applications
- Spring Boot accelerates development of microservices by automating routine tasks
- React JavaScript offers amazing modularity with component-based structure
- The resilience of this system is fortified through technologies like Kubernetes and Docker to help contain, maintain and sustain the unpredictable demand surges.

Future working would include a fully implemented CI/CD pipeline to rapidly improve deployment time and iterative updating.

Emotion Recognition Using Multiple Physiological Signals

Sherin Khalil, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

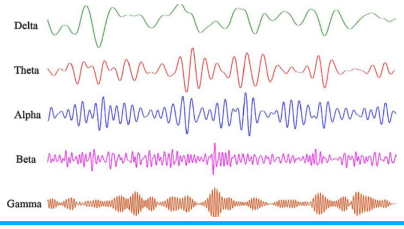
Emotions significantly influence human behaviour and decision-making, particularly in a digital era dominated by human-computer interactions. They can be expressed in various forms. We defined emotional states in three different ways according to the values of valence, arousal, dominance provided in the data set. The main objective of this project is to comparatively analyse the performance of various machine learning (ML) classifiers to accurately and speedily detect human emotional states using multiple physiological signals. This study is conducted with a Dreamer dataset based on our defined emotional states. This dataset comprises of both raw EEG and ECG signals. We found that ensemble is the best ML classifier for binary emotion detection is ensemble learner, the best classifier for non-binary emotion detection is multinomial logistic classifier, when number of discrete emotional states increases, the accuracy of the classifiers decreases. This study makes an impact on monitoring patients' mental healthcare conditions.

Research Questions

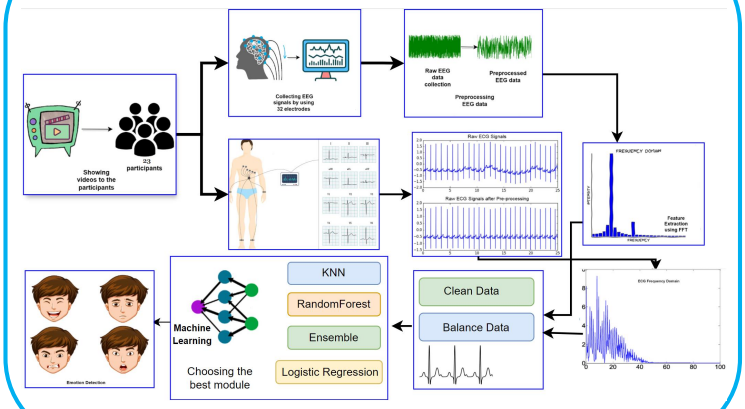
- ❑How effective are machine learning classifiers in predicting emotional states based on physiological signal?
- ❑What is the relation between form emotion states and the performance of classifiers?

Data Preprocessing

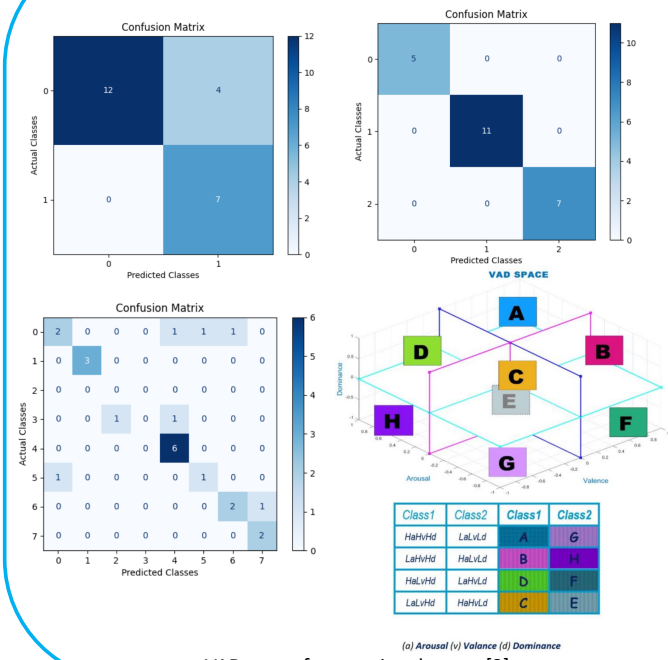
- ❑ Raw data were collected from 23 participants watching 18 videos.
  - The EEG signals were collected 14 electrodes.
  - The ECG signals were collected 2 electrodes right and left heart.
- ❑ The EEG signals were extracted and processed with FIR filters.
- ❑ All outliers and infinities were identified and removed.



System Architecture



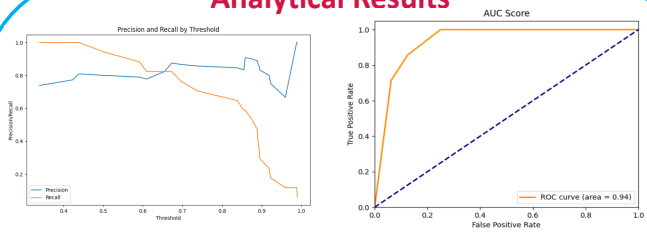
Emotional States



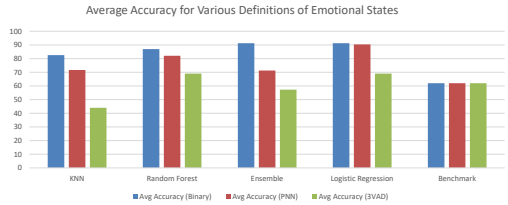
(a) Arousal (v) Valance (d) Dominance

VAD space for emotional states [2]

Analytical Results



Methods	KNN	Random Forest	Ensemble	Multinomial Logistic Regression	Benchmark
AVG Accuracy	82.61%	86.96%	91.30%	91.30%	≈62%
AVG Precision	63.64%	70%	77.78%	8.70%	NA
AVG Recall	0.75%	81.25%	87.50%	93.75%	NA
AVG AUC	87.5%	90.7%	93.75%	100%	NA
AVG Rate Error	17.39%	13.04%	8.70%	77.78%	NA
AVG Specificities	75%	81.25%	87.50%	87.50%	NA
AVG F1-Score	77.78%	82.35%	87.50	87..50%	≈59%



Data Balancing and Classifiers Validations

- Before Balancing 207 training set and 23 test set; After Balancing 240 training set and 23 test set
- Balancer – SMOTE + Near Miss version 1
- Classifiers – 5NN, RandomForest, Ensemble, Logistic Regression in Multinomial kernel
- Cross validation – Repeated 10-fold cross validation

Conclusions

- Both ensemble and logistic regression classifiers strike higher accuracy 91.3% for binary emotional states, its accuracy outperforms the accuracy in the benchmark.
- For the non-binary emotional states, both random forest and logistic regression tend to have a higher accuracy; for the PNN emotional states, both classifiers strike to around 90%; for the 3D VAD space, it only managed to get around 69% accuracy. However, it is 7% higher than the benchmark accuracy.

References

1. Katsigiannis, Stamos and Ramazan, Naeem, 98--107, DREAMER, *2017*.
2. Umran Isik, Aysegul Guven and Turgay Batbat, 2141, 3D VAD.

Acknowledgments

I would like to extend my sincere thanks to my supervisor, Jing Hua Ye, for all the help provided, and to Ruairi O'Reilly for all the support.

Movie Recommendation System

Mark Lawton, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

1. Introduction

The aim of this final year project is to assist people in making informed movie choices. Choosing the perfect movie to watch can often feel like searching for a needle in a haystack. This project will simplify that search, with a custom-built web scraper to gather data from IMDB and Rotten Tomatoes, this data is then fed into the recommendation algorithm that generates a list of 10 recommended movies. Which are presented on the website.

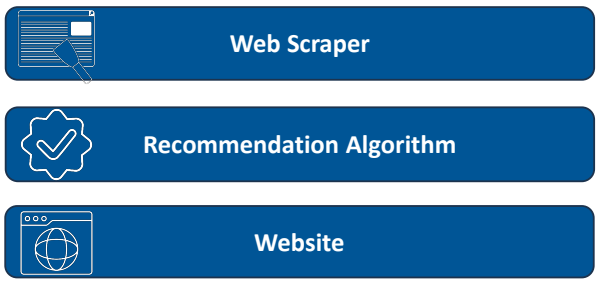
2. Objectives

1. To study **Webscraping** for facilitating the collection of movie information from IMDB and Rotten Tomatoes such as, genre, cast and director, to enable the creation of a database for the recommendation algorithm to make suggestions to the user.

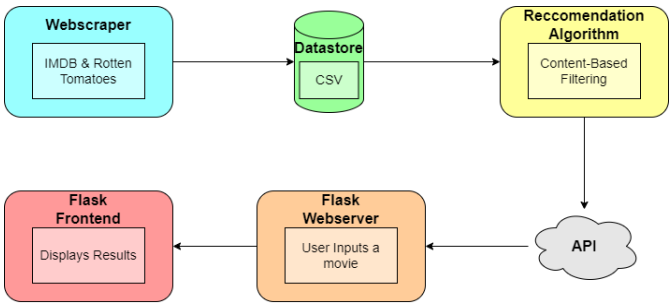
2. To study **Recommendation Algorithms,** so that the website would suggest a list of 10 movies in descending order from most recommended to least recommended.

3. Design a **Website** to provide a friendly and aesthetically pleasing way for user interaction.

3. Core Components



4. Architecture



5. Ethical Concerns

To follow ethical data collection the following measures were implemented.

- **Scheduled Scraping –** The system is programmed to run semi-annually, to reduce the frequency of requests to IMDB and Rotten Tomatoes servers.
- **Limited Dataset –** The quantity of movies is limited to 1000 to ensure a less intrusive retrieval process.
- **Essential Data –** The scraper targets only vital data, focusing on ratings and other core movie details.

6. Technology

7. Results

Movie Title	Cast	Genre	Plot Synopsis	IMDB Score	Director	Year	Formatted Title	Audience Score	Critic Score
Everythin	['Michelle	['Action', 'Ad	A middle-aged Ch	7.8	Daniel Kwar	2022	everything everywh	94%	86%
The King	['Tom Glyn	['Biography',	Young Henry V er	7.3	David Michc	2019	the king	71%	84%
How to Tr	['Jay Baruc	['Animation'	A hapless young	8.1	Dean DeBlo	2010	how to train your dr	99%	91%
Wonka	['Timothée	['Adventure'	With dreams of o	7	Paul King	2023	wonka	82%	91%
Ghostbus	['Bill Murra	['Action', 'Co	Three parapsychc	7.8	Ivan Reitma	1984	ghostbusters	95%	88%
Godzilla M	['Minami H	['Action', 'Ad	Post war Japan is	8.3	Takashi Yam	2023	godzilla minus one	98%	98%
Sing 2	['Matthew	['Animation'	Buster Moon and	7.3	Garth Jenni	2021	sing 2	72%	98%
Society of	['Enzo Vog	['Adventure'	The flight of a rug	7.8	J.A. Bayona	2023	society of the snow	90%	87%
Cars	['Owen Wi	['Animation'	On the way to the	7.2	John Lasset	2006	cars	75%	80%
BlackBerr	['Jay Baruc	['Biography',	The story of the r	7.4	Matt Johnsc	2023	blackberry	98%	93%
The Incre	['Craig T. N	['Animation'	While trying to le	8	Brad Bird	2004	the incredibles	97%	75%
Tangled	['Mandy M	['Animation'	The magically lon	7.7	Nathan Grei	2010	tangled	89%	87%
Puss in Bc	['Antonio E	['Animation'	When Puss in Bo	7.8	Joel Crawfo	2022	puss in boots the las	95%	94%

```
Movie Entered - The Dark Knight
Rotten Tomatoes
1. The Batman
2. Inception
3. Interstellar
4. Godzilla vs. Kong
5. Kingdom of Heaven
6. The Super Mario Bros. Movie
7. Shrek
8. The Bad Guys
9. The Lion King
10. Gladiator
```

```
Movie Entered - The Dark Knight
IMDB
1. The Batman
2. Interstellar
3. Inception
4. Kingdom of Heaven
5. Godzilla vs. Kong
6. The Super Mario Bros. Movie
7. Shrek
8. The Bad Guys
9. The Lion King
10. Gladiator
```

Factoring in age rating would help get more accurate recommendations.

8. Acknowledgments

I would like to thank both my first semester supervisor Helen Fagan and my second semester supervisor Arthur Tobin for their help and guidance throughout this project.

Towards Open-Source
Refactoring and Migrating Legacy Systems as OS Solutions

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Wei Xuan Leong, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction & Objective

Embracing the open-source model can foster values such as community collaboration, software transparency and sustainable development effort [1]. However, legacy application to open-source migration entails challenges such as decoupling proprietary dependencies and ensuring thorough documentation, all while maintaining functionality. In this project, a pragmatic example will be undertaken in the form of refactoring and migrating PEM, a pre-existing application to an open-source model.

Challenges

- **Code coupling & Dependencies:** Potential to restrict flexibility, hinder scalability, and increase the risk of software maintenance issues.
- **Inconsistent Database Schema:** Inconsistent database schemas from diverse contributions risk application stability, highlighting the need for standardized structures.
- **CI/CD Workflows:** Contributors may encounter challenges in handling code changes, testing, and deploying new application versions.
- **Lack of Documentation:** Lack of OS documentation hinders maintenance, development, and community engagement, impeding software progress and project credibility.

Architectural Change

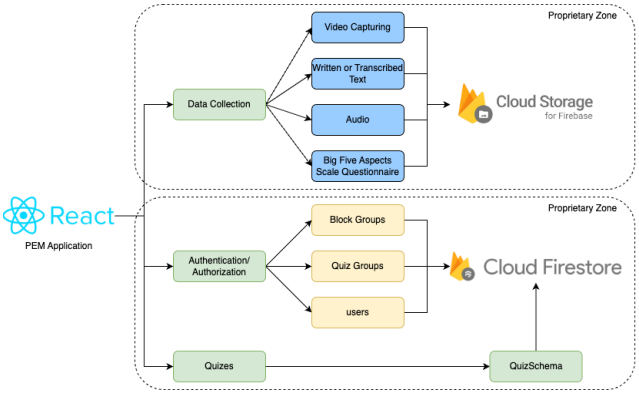


Figure 1: Original PEM Architecture

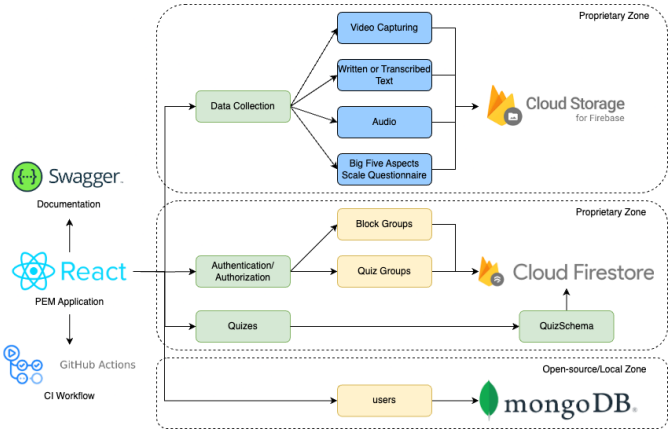


Figure 2: Final PEM Architecture

Results & Conclusion

Decoupling dependencies: Decoupling of proprietary software such as Firebase to open-source options like MongoDB. This provide a more maintainable and localizable application.

Separation of Concerns: Front-end and Back-end: Incorporating best practices such as Don't Repeat Yourself (DRY) in optimizing the backend API code and reducing unnecessary HTTP requests through factorizing fetch queries in the frontend. This approach increases cohesion and efficiency, enhanced the overall performance of the system.

Continuous Integration Workflows: Leveraging GitHub actions to implement formatting checks whenever a commit is pushed to the repository ensuring consistent formatting throughout the entire codebase anytime. This practice minimizes inconsistencies for developers, promoting stability and efficiency in development tasks.

Documentation: Implementation of a comprehensive README file that offers essential information about PEM and API endpoint documentation and testing application using SwaggerUI. These are crucial for maintaining a sustainable and detailed OS application repository.

References

1. "What is open source?" [Online]. Available: https://www.redhat.com/en/topics/open-source/what-is-open-source

Acknowledgments

The author would like to thank Dr Ruairi O'Reilly for supervising this project and Adam Baldwin and Ryan Donovan for their knowledge and guidance in this area of study.

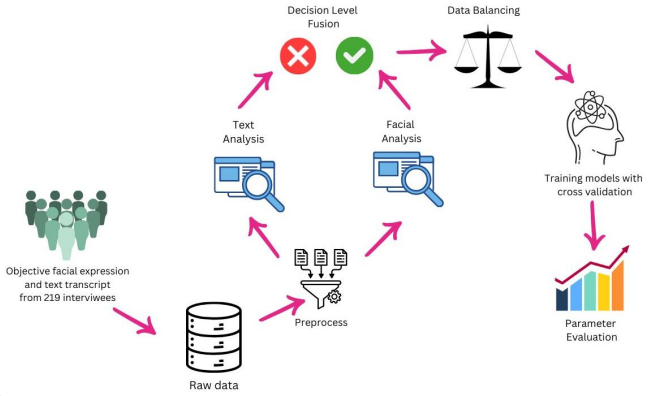
Depression Detection With Intelligent Emotion Models

Meng Jie Liu,
BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

Depression detection is one of biggest challenges in mental healthcare field using AI models, most studies are mainly focus on single modality or various forms of multimodalities. Our study aims to comparatively investigate the performance of various machine learning classifiers and deep learning classifiers for accurately identifying the binary view of depression. Such a comparison study has been conducted with the Extended Distress Analysis Interview Corpus (E-DAIC) dataset, which comprising both the transcripts of interviewees and preprocessed facial expressions. All these information are fused with the decision-fusing technique. This analytical study revealed that the deep learning classifiers, like CNN, outperform the machine learning classifiers. In terms of machine learning classifiers, our study indicated that there are no significant performance differences among machine learning classifiers.

System Architecture



Research Questions

- Is there a significant difference in the classification of depression when applying oversampling versus undersampling methods to balance the dataset?
- How effective are multimodal data sources in improving the accuracy of depression detection compared to unimodal data sources?
- How does the fusion of decision-level outcomes from various classifiers affect the reliability and accuracy of depression diagnosis?

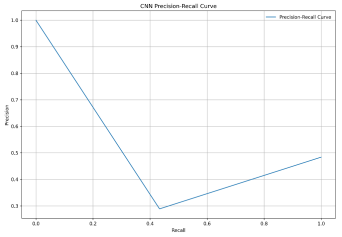
Data preprocessing

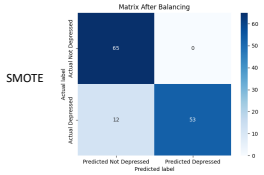
- Transcript Data Preprocessing - Stopword removal, lemmatization, and TF-IDF vectorization.
- Facial Expression Preprocessing – The original data was preprocessed with OpenFace application, therefore, we applied the rigid data cleaning, like removing blank space and remove any non-numerical values.
- Data Balancing Techniques
  - Oversampling - SMOTE
  - Undersampling – Random undersampling
  - Hbrid - SMOTE + Tomek

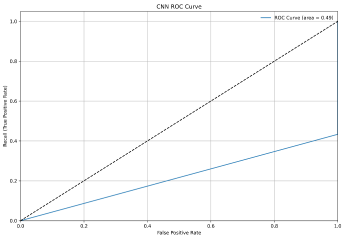
Classifiers Training and Validation

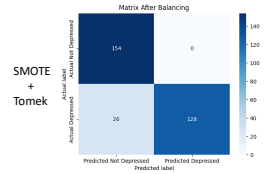
- Machine Learning Classifiers - XGBoost, Random Forest, Logistic Regression
- Deep Learning Classifiers - Convolutional Neural Networks (CNN)
- Train and Test – Training set 175 (80%), Test set 44 (20%)
- Cross-Validation - The use of K-Fold cross-validation to assess the robustness of the models

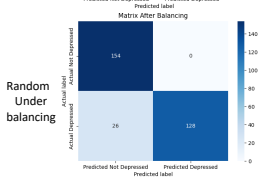
Analytical Result











Model	AVG Accuracy	AVG Precision	AVG Recall	AVG F1	AVG False Positive Rate	AVG Specificity	AVG AUC	AVG Error Rate
XGBoost	79%	85%	78%	77%	0	1	78%	20%
Random Forest	79%	85%	78%	77%	0	1	78%	20%
Logistic Regression	79%	85%	78%	77%	0	1	78%	20%
CNN	91%	100%	82%	89%	0	1	82%	8%
BenchMark[1]	90.33%	91%	91%	91%	/	/	/	/

Conclusions

- The XGBoost, Random Forest, and Logistic Regression models show consistent performance across all metrics. With an average accuracy, precision, recall, and F1-score all in the range of 77% to 85%, it indicates that no single model significantly outperformed the others on the given dataset for these algorithms.
- The CNN model stands out with the highest average accuracy of 91% indicating its robustness in correctly identifying depression when present.
- Compared with the benchmark model, the CNN outperforms in average accuracy and precision but falls slightly short in recall and matches in F1-score. This comparison suggests that while the CNN is more cautious in predicting depression, it may miss some cases that the benchmark model would catch.

References

1. Bhargava, Chaitanya. "Depression detection using sentiment analysis of tweets." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12.11 (2021): 5411-5418.

Acknowledgments

The author would like to thank Jing Hua Ye, Ruairi O'reilly for their supervision.

Beyond the Hook: A Phishing Awareness Campaign

Kevin McCarthy, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024

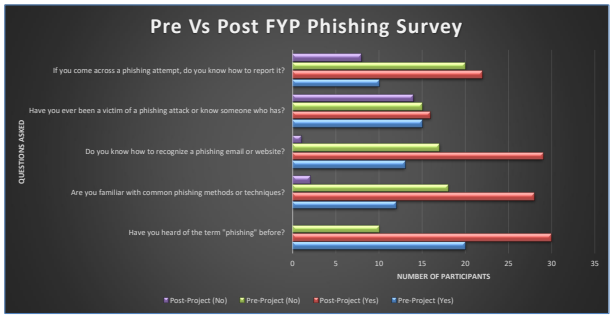
Introduction

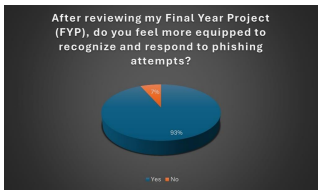
Phishing continues to be an ongoing threat to both individuals and organisations all over the world. This project intends to create novel ways to effectively prevent phishing, improve user education and promote awareness in response to this urgent cybersecurity threat. This project aims to equip users with the knowledge and resources needed to identify and reduce risks posed by phishing using a combination of research, technology development and user interaction. Through tackling the technological and human aspects of phishing, this study adds to the continuous endeavours to establish a more secure online environment for everyone.

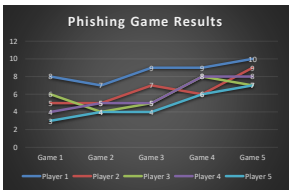
Amygdala Hijack

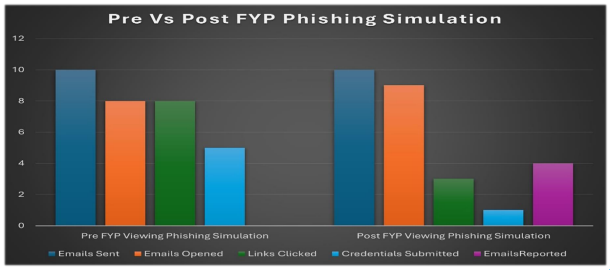
The **frontal cortex** governs decision-making and rational thinking, while the **amygdala** processes emotions and threat responses. Phishing exploits the 'amygdala hijack,' triggering immediate emotional reactions that override rational judgment. This neurological shortcut leads to impulsive actions, such as clicking on malicious links or sharing sensitive information, before the frontal cortex can analyse the situation. Understanding this interplay between brain regions is crucial for combating phishing effectively.

Research Findings



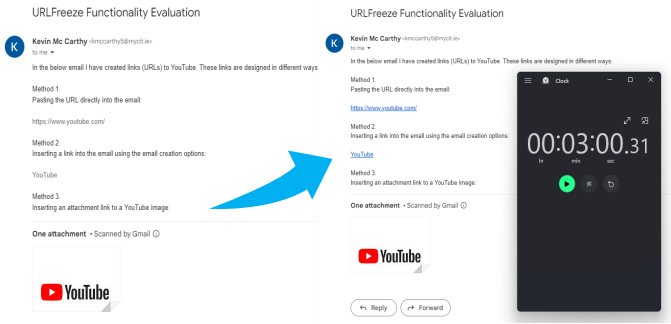






Project Development Highlights

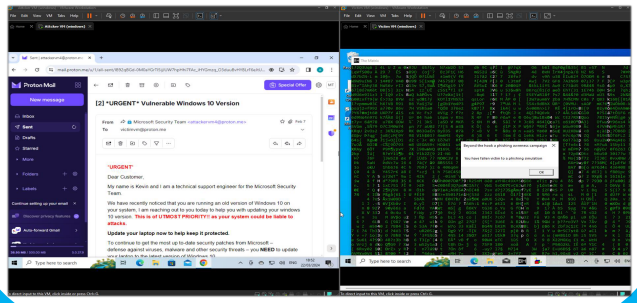
Browser extension demo showcasing link deactivation on the left and reactivation after 3 minutes on the right after opening an email.



During the project I created a phishing demonstration to highlight techniques and consequences:

Phishing Email Sent

Consequences of Phishing Attack



Acknowledgments

The author would like to acknowledge the support and guidance of Gerard Mac Sweeney.

Conclusions

- I investigated the intricacies of phishing tactics and the psychological underpinnings of them, such as the phenomenon of amygdala hijacking.
- To counter these dangers, I created a comprehensive solution that included:
  - A browser extension designed to lessen the impact of the amygdala takeover by disabling links within an email for a time period.
  - An educational website hosting a phishing demonstration video, URL checker and a phishing game.

Message Routing Transformation Engine

Kacper Michalak, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, April 2024

Introduction

The purpose of the Message Routing Transformation Engine is transformation and delivery of financial messages originating from JCCS to other systems. The current implementation is many years old and lacks extensibility. The goal of the project was to redesign the code base and introduce extensible components, to keep the existing functionality but also add a new Message In/Queue Out system. The project also added an interactive dashboard to provide visibility on message transformations including statistics.

Key Features

1. File Upload and Queue System
2. Data Validation
3. Conversion and Storage
4. View, Filter and Search Functionality
5. Statistical Insights
6. User-friendly Interface

Main Technologies Used

1. Java
2. Spring Boot
3. Thymeleaf
4. Apache Camel
5. Apache ActiveMQ
6. Java Persistence API (JPA)

Spring Boot + Apache Camel

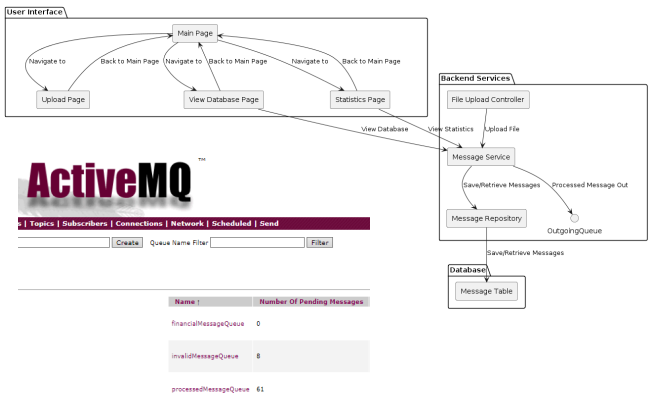
Java

Financial Message Challenges

1. Data Integrity
2. Message Formats
3. Security
4. Real-Time Processing
5. Error Handling
6. Regulatory Compliance

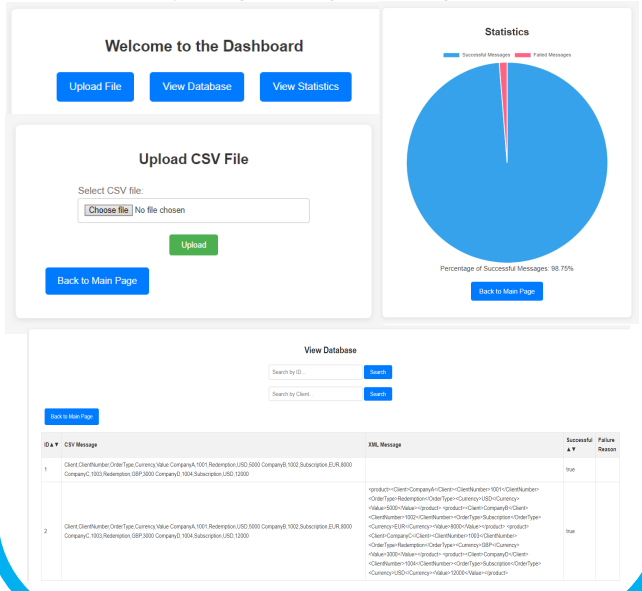
Queue System Implementation

A pivotal aspect of the project involved the implementation of a sophisticated queue system to manage the flow of messages efficiently and reliably. Drawing upon industry best practices, the queue system was made to address the unique requirements and challenges of the financial messaging domain. It added a new Message In/Queue Out functionality.



File Upload and Dashboard Functionality

The development of the File Upload and Dashboard functionality was a fundamental component of the project. It was aimed at providing users with a streamlined and intuitive interface for managing financial messages effectively and gain insights through statistics.



Advantages of New Design

The new design offers improved adaptability through streamlined message routing and transformation. It enhances extensibility, enabling seamless integration of new message formats and protocols. Additionally, it provides a user-friendly interface for efficient file upload, message processing, and visualisation of key statistics, enhancing overall system efficiency and provides the user with an easy-to-use interface.

Conclusions

The project represents an advancement in the realm of financial messaging systems, addressing key challenges faced by the existing system while introducing innovative features to enhance productivity, adaptability and user practical interface.

Acknowledgments

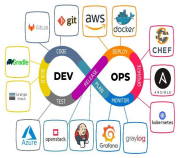
Helen Fagan – Term 1 and 2 supervisor
Robert Scully – Vestima Prime Application Team Manager

Optimizing DevOps Practices for Cloud-Native Application Development

Jason Molina, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024

Introduction

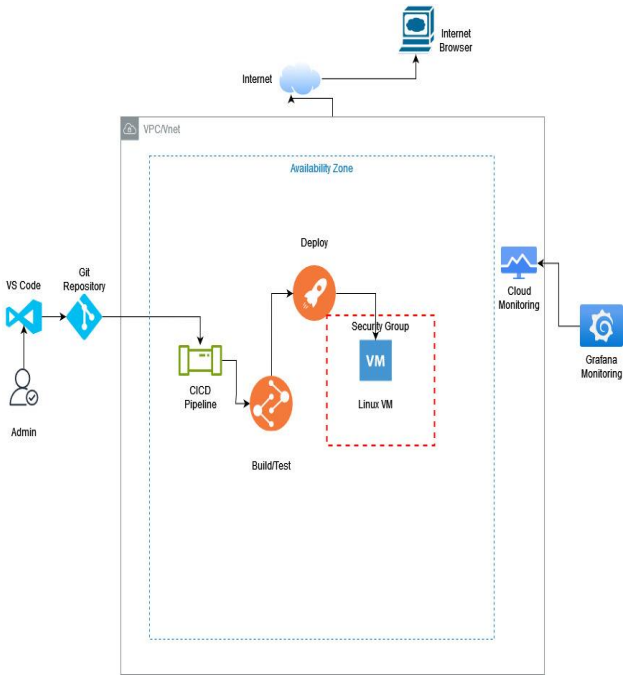
In today's fast-paced digital landscape, integrating DevOps practices with cloud-native application development is essential for organizations aiming to innovate and deliver value efficiently. "Optimizing DevOps Practices for Cloud-Native Application Development" explores DevOps tooling across three leading cloud platforms: Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP).



Problem Statement & Solution

By implementing and testing CI/CD pipelines based on specific criteria, the project provides an in-depth assessment of functionality, scalability, and reliability. The findings aim to identify the strengths and weaknesses of each platform's DevOps offerings, helping organizations make informed decisions for their DevOps initiatives. This project bridges the gap between theory and practice, offering actionable insights to navigate DevOps implementation across AWS, Azure, and GCP effectively.

Architecture



Project Goals

- Conduct a comprehensive comparative analysis of DevOps tooling across AWS, Azure, and GCP.
- Evaluate the capabilities and features of Continuous Integration and Continuous Deployment (CI/CD) pipelines on each platform.
- Assess the performance, scalability, and reliability of DevOps tools and services offered by each cloud platform.
- Implement CI/CD pipelines on AWS, Azure, and GCP to simulate real-world deployment scenarios.
- Test the automation capabilities, deployment flexibility, and monitoring tools of each platform's DevOps offerings.

Project Deliverables

- Identify the strengths and weaknesses of each cloud platform's DevOps tools based on predefined criteria.
- Provide actionable insights and recommendations to help organizations select the most suitable cloud platform for their DevOps initiatives.
- Analyse the cost-effectiveness and ROI (Return on Investment) of implementing DevOps practices on each platform.
- Explore the broader implications of adopting DevOps practices within the context of cloud-native application development.
- Empower organizations with the knowledge and tools needed to streamline development workflows, optimize DevOps practices, and accelerate cloud-native application delivery effectively.

Tech Stack

aws Azure Google Cloud Platform Visual Studio python Flask GitHub Grafana

Acknowledgments

I would like to thank my supervisors in semester 1 & 2, Dr. George O' Mahoney and Dr. Pat McCarthy for their support and guidance throughout this year.

Mobile Application to Improve Speech

MTU Ollscoil Teicneolaíochta na Mumhan Munster Technological University

Daniel Murphy, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2022

Introduction

This is an Android application capable of recording your speech and comparing this to a generated sample recording of the text to generate a comparison score.

This app allow the user to practive thei speech in a relexed enviroment and gain feedback on their progress

Implentation

In this project there is a number of metrics which I will be evaluation my project. As the application delivers a score result, evaluating the application's effectiveness is made easier through this defined metric. There is also non-functional requirements which also require there own metric to evaluate them

The term 'Fingerprint' is used with relation to the decoded audio file. Within the WavAudioCompare class function, two values are produced. The Similarity value is measured from 0~1, which 0 means no similar feature is found and 1 means in average there is a least one match in every frame. The score value measures the similarity score of the fingerprints in regards to the number of features found in the fingerprints per frame.

This metric determines if the text entered is able to be spoken by the Text to Speech engine and save the file for later use.

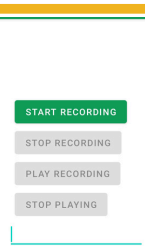
In order to test the similarity between the two audio files, the user must first enter the text they wish to use as a comparison and then press the record button. The user then records themselves speaking the text. Once the user hits stop record, the recording is compared to the Text to Speech Output. The Similarity and the Score are then displayed on the screen and saved to a text file

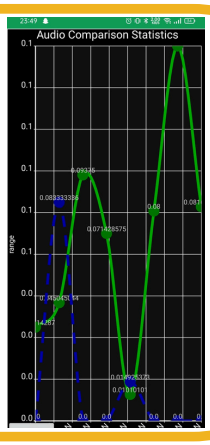
In order to better communicate the results of the comparison, the statistics are loaded into a visual graph. Testing with a combination of good and bad results, the graph visually indicates this with a line going up and down through he values in order

Conclusion

feel I accomplished I accomplished the main goals I had set out for this app. I made a functional Android application free of any major bugs and errors. The application is able to Record Audio successfully and Compare it to Speech generated from user inputted text, with he results printed on a graph. The hope is this application could be used by those hoping to improve their speaking ability, giving a reliable training tool for those public speaking or trying to improve their reading comprehension

Screenshots





Acknowledgements

Thank you to Alison o Shea and Brian Murphy for their support

Approaches to Visualising Security Monitoring Data

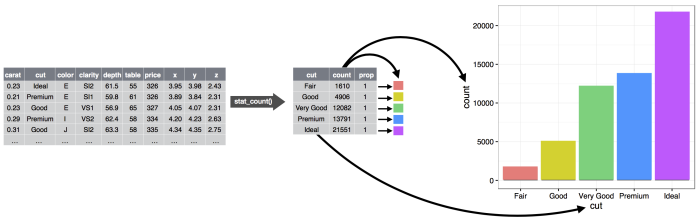
Shauna Murphy, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024

Introduction

The aim of this project is to investigate six data visualisation platforms and compile an in-depth comparative analysis. Case studies for the education and healthcare sectors were undertaken with the aim of suggesting specific visualisation platforms most suitable to each sector. Given that data visualisation is lacking, a lot of security monitoring data is not being utilised to its fullest potential. Large organisations are producing large amounts of raw data quickly; this phenomenon is referred to as "big data." Big data involves processing a lot of different types of data quickly at a high volume, which makes it challenging to keep up with at such a rapid pace of change.

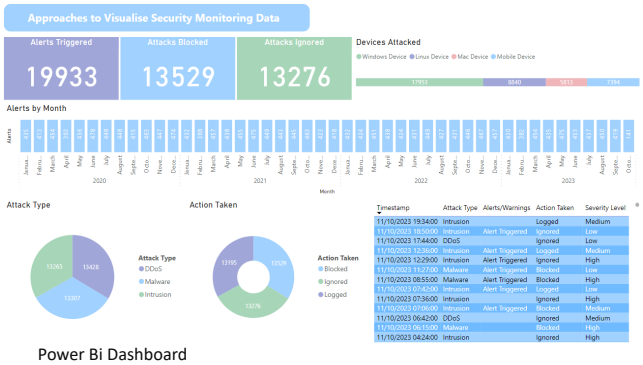
Objectives

- Visualise data sets
- Investigate open-source visualisation tools
- Analyse and compare different data visualisation platforms

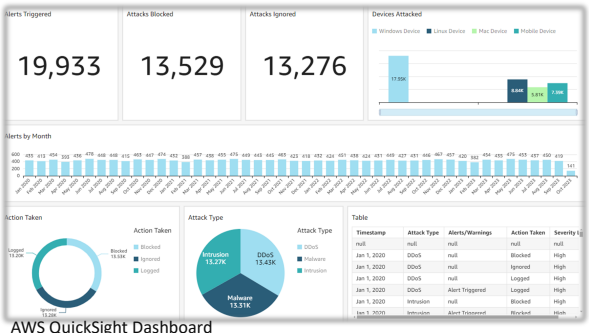


Data Visualisation Dashboards

Security monitoring visualisation dashboards were created on Power Bi, Tableau, AWS QuickSight, ServiceNow, Splunk and Grafana.



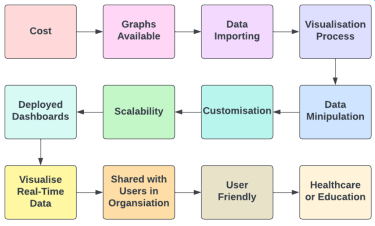
Power Bi Dashboard



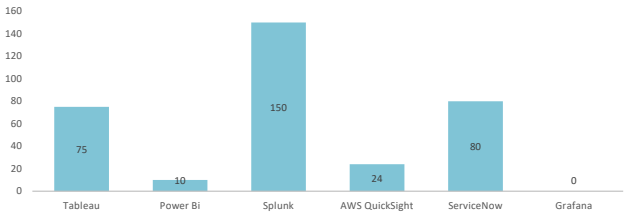
AWS QuickSight Dashboard

Comparison

The comparative analysis involved comparing the six platforms in terms of 12 features, including cost, graphs available, data importing, scalability and real-time data.



Cost per Month



The cost of the platform will play a big role in the organisation choosing a platform to use, as well as the scalability of the platform and the option to visualise real-time data. Case studies for the healthcare and educational sectors were conducted in order to identify which platforms are most suited to the specific requirements of each.

Conclusions

- There is a need for data visualisation and more learning resources.
- Organisations are moving from Tableau to Power Bi.
- Although primarily used as a management system for IT services and assets, ServiceNow has many more uses, including data visualisation.

Future Work

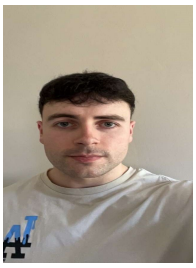
- Analyse further data visualisation platforms
- Develop a set of step-by-step guidelines for data visualisation
- Expand the analysis to incorporate further data sets.

Acknowledgments

The author would like to acknowledge the support of her supervisor David Murphy for all his guidance during the implementation of this project as well as the MTU computer science department.

An Analysis of Start-up Security & Development of a Cost-Effective Tool

Adam Murray, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024



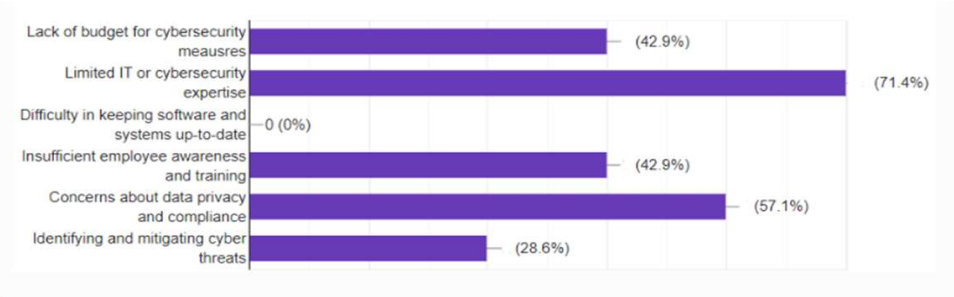
Introduction

Start-up companies are promising and often the life of a country's economy, however they are often constraint due to their limited resources, especially in cybersecurity.
In this project, I carry out a deep analysis of what start-up and small to medium business struggle with, in terms of cybersecurity and provide cost effective solutions to become more secure.

Analysis Findings

I carried out multiple in person and online questionnaires with start-ups in Cork, on cybersecurity. Here are my main findings:

- **Lack of budget for cybersecurity.**
- **Limited cybersecurity expertise.**
- **Insufficient employee awareness and training.**
- **No understanding on how to protect against cyberattacks.**



StartUp Defence Solution

From my research and discussing with industry experts, my project focused on the following:

- An intrusion detection system using Snort to cover the most common threats faced by start-up companies such as Nmap scans, brute force attacks, and DDoS attacks.
- Utilization of Splunk to provide field extractions for Snort alert logs.
- A web application that educates users on the most common threats faced by start-ups. The web application contains overviews, mitigations, and demonstrations of attacks.

Technologies Used

splunk> SNORT WiX

Acknowledgments

I would like to acknowledge the support of Mary Davin for advice and guidance throughout this project, members of MTU, and work colleagues for their continuous support.

Sheldifier: Indoor Air Quality Management System

Shelton Ngwenya, BSc Honours in Software Development
Department of Computer Science, MTU Cork, May 2024

Introduction and Motivation

- Enhancing IAQ monitoring in educational environments for improved cognitive function and health.
- Driven by the need for a responsive ventilation system that utilizes real-time IAQ data to maintain optimal conditions.

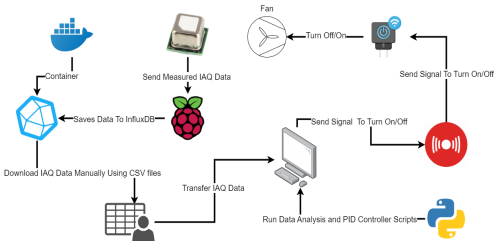


Figure 1. System Architecture

Perceived vs. Measured Air Quality

- Measured with sensors (CO2, temperature, humidity) indicating the physical conditions.
- Perceived quality assessed through surveys, focusing on comfort, freshness, and cognitive impact.
- Logistic regression models to correlate perceived and measured quality.

Figure 2. Survey for POE Perceived IAQ

	Question	Possible answers
1	Based on how you felt in the room today, rate your comfort level related to the air's moisture content. *Aims to answer: This is aimed to understand the occupants' comfort regarding humidity levels. It helps correlate subjective perceptions of moisture content with relative humidity data from the sensors. High or low extremes can affect comfort, and health within the room.*	1: Very Uncomfortable 2: Uncomfortable 3: Neutral 4: Comfortable 5: Very Comfortable
2	Rate your level of alertness throughout the day in this room. *Aims to answer: This seeks to understand the cognitive effects of air quality, particularly how CO2 levels and ventilation might influence occupants' alertness and productivity. Elevated CO2 levels can decrease cognitive function and alertness, so this question helps link*	1: Completely Unalert 2: Somewhat Unalert 3: Neutral 4: Alert 5: Extremely Alert
3	Did you experience discomfort due to the temperature in the room today(hot or cold)? *Aims to answer: This question assesses thermal comfort. Responses can help us correlate subjective discomfort with actual temperature data collected by our sensors. It directly addresses whether the temperature settings in the room meet the occupants' comfort levels or if adjustments*	Yes. No.
4	During your time in the room today, how would you rate the freshness of the air on a scale from 1 *Aims to answer: This targets the perception of air freshness, indirectly related to CO2 concentration and ventilation effectiveness. It's crucial for understanding if the air exchange is sufficient to maintain a feeling of freshness.*	1: Very Stale 2: Somewhat Stale 3: Neutral 4: Somewhat Fresh 5: Very Fresh
	Any Additional Comments:	

Predictive Analytics & Control

- ARIMA models predict IAQ based on historical sensor data.
- Multivariate regression analysis relates measured IAQ parameters to perceived air quality scores.
- FLC interprets model outputs for nuanced ventilation control.
- PID control algorithms dynamically adjust system settings to maintain desired IAQ levels.

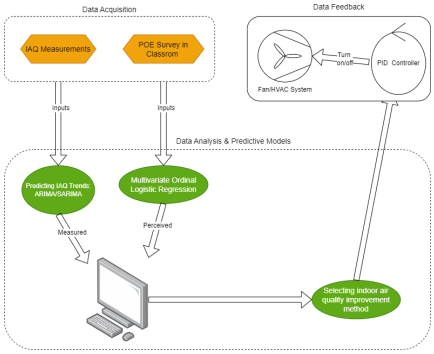


Figure 3. System Flow Diagram

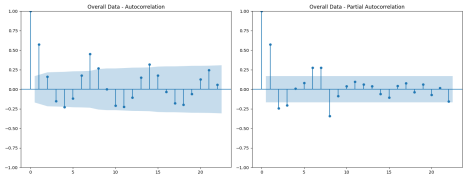


Figure 4. CO2 ACF vs PACF Diagram

System Architecture & Data Insights

- Built on Raspberry Pi and sensors with InfluxDB for data storage and python scripts for data analysis.
- Analysis of data reveals auto-correlation, trends, and patterns in IAQ, enabling predictive management.
- Continuous feedback loop ensures system adaptability to changing environmental conditions.

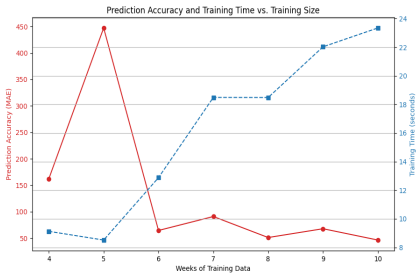


Figure 5. Model Fitting Performance Diagram

Impact & Contributions

- Provides a comprehensive understanding of both perceived and measured IAQ.
- Demonstrates the efficacy of integrating IoT technology for real-time IAQ monitoring and management.
- Serves as a pioneering model for IAQ management in educational and other public settings.

Acknowledgements:

Special thanks to:
- Victor Cionca (supervisor).
- Charmaine Ngwenya (sister).

for guidance and support

Glossary

1. IAQ = Indoor Air Quality
2. CO2 = Carbon dioxide
3. POE = Post Occupancy Evaluation
4. PID = Proportional-Integral-Derivative
5. FLC = Fuzzy Logic Controller
6. IoT = Internet of Things

Network Security Awareness

Creating an educational packet sniffer

Rory O'Callaghan, BSc Honours in IT Management
Department of Computer Science,
MTU Cork, May 2024

Introduction

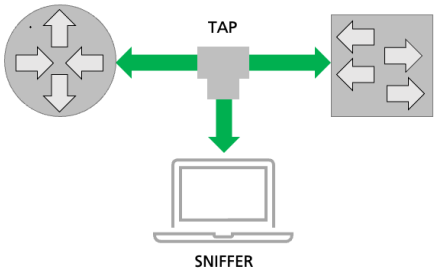
- Many new students face difficulty when entering the world of networking and have a hard time learning new concepts and buzzwords commonly used in the industry.
- There is a lack of accurate and concise learning resources for existing sniffer applications like Wireshark, so it's only right that there is a tool that can meet the educative and functional needs of these students.

Problem Solution

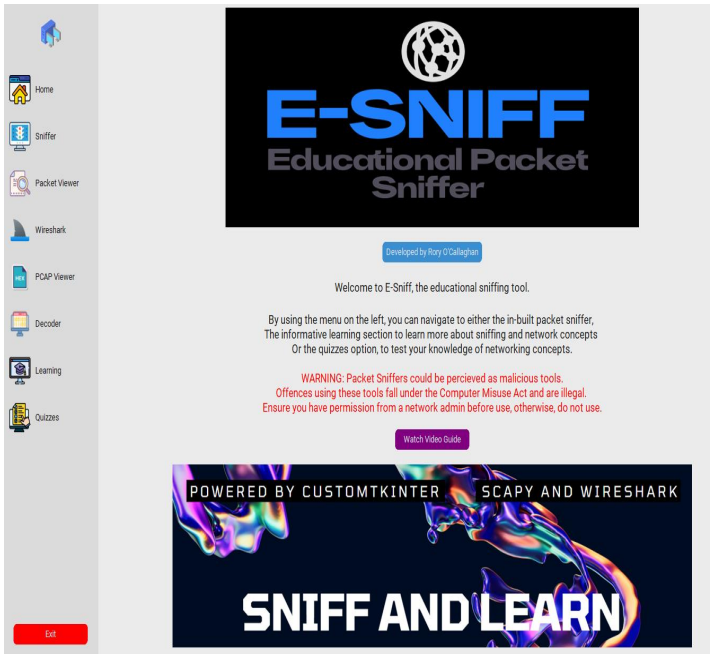
- This project educates the user using learning resources while providing the same sniffing functionality as other applications on the market to assist any new user on its usage.
- The E-Sniff application comes with a range of features like a Scapy packet sniffer, packet viewer, Wireshark sniffer, PCAP viewer, packet decoder and a range of learning materials paired with informative quizzes that allow the user to track their learning and integrate themselves into the world of networking.

What is Packet Sniffing?

- Sniffing, whether malicious or for security, is the act of scanning network packets or messages between devices, users or apps across a network.
- A user who wishes to keep their network secure would seek malicious or shady looking packets on their network and attempt to stop them, while a threat actor may see a compromised network through sniffing and attempt any number of malicious attacks.



E-Sniff Application



Technologies Used

Wireshark, scapy, SQLite, CustomTkinter

Acknowledgments

The author would like to acknowledge Mary Davin for supervising this project and all IT Management lecturers for their expertise and concise teaching to help me understand these areas of study. This project would not have been possible without their consistent assistance.

Exploring the Function of Keyloggers in Credential Theft

Seán O'Connor

Software Development - level 8 - Munster Technological University

Abstract

This projects aim is to explore the world of Keyloggers and how they are utilized in credential theft while also raising awareness about digital security. Part of this project involves writing a basic software-based Keylogger in Python and setting up a secure virtual environment to facilitate the safe execution of the malware to observe its functionality and behaviour. This project will provide insight into the state of the art of Keylogger generation, delivery and functionality and will explore the state of the art in detection and prevention of said Keylogger.

Research objectives

1. Educational awareness of digital security
2. Demonstration of Keyloggers and their functionality
3. Highlight Security Risks
4. Suggestion of preventative measures

Proposed Approach

A scrum approach was used to manage this project, involving 4 sprints, each with a duration of 3 weeks and their own tasks to be completed within this duration.

- Sprint 1: Configuring secure lab environment and commencing development of Keylogger script.
- Sprint 2: Continuing development and testing of Keylogger functionality.
- Sprint 3: Carrying out simulated attacks within the secure lab environment.
- Sprint 4: Create video demonstration of the Keylogger in action and suggest preventative measures against such attacks.

Proposed Architecture

Below is a diagram illustrating the architecture of the system. A Windows 11 laptop was used as the host machine, using VirtualBox as the hypervisor to host two virtual machines which were isolated from the host and the internet. Both machines were connected to a NAT network to allow communication with one another. This system replicated two separate machines, one being the target machine used by the victim, and the other being the attacker machine used by the threat actor. This is is the system that was used for demonstrating the process of a cyber attack and the execution of the malware without infecting the host machine.

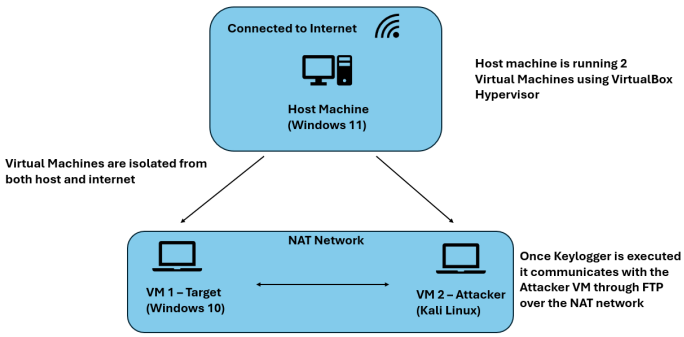


Figure 1. Architecture Diagram

Testing and results

During the development of the software based keylogger, various functionality was considered such as:

- Execute Keylogger on start-up through a scheduled task
- Record all keystrokes inputted through keyboard
- Encrypting recorded keystrokes
- periodically send recorded keystrokes to attacker machine through FTP

These features allow the Keylogger to covertly record the input from an average user without their knowledge. While testing these features, the Keylogger successfully captured keystrokes and stored the encrypted data in a log file hidden on the machine ready for exfiltration through FTP.

Technologies Used

References

1. "Keylogger Detection: A Systematic Review" by Ekele Victoria, Adebiyi Ayodele, Igbekele Emmanuel.
2. "Design, Analysis and Implementation of an Advanced Keylogger to Defend Cyber Threats" by Bejo, Sahil Prasad and Kumar, Biresh and Banerjee, Pallab and Jha, Pooja and Singh, Amar Nath and Dehury, Mohan Kumar.

Real World Examples

HawkEye Remote Access Trojan

- 19 March 2020, phishing emails masquerading as legitimate emails from the General Director of the World health Organisation.
- Emails contained an .EXE for a cure for the Covid-19 virus.
- The .EXE contained the malware know as HawkEye. This malware had anti-VM and anti-sandbox capabilities and disabled Windows Defender via the registry and disabled scans and updates via Powershell.
- It collected credentials from keystrokes, and web browsers. it then encrypted the collected data and exfiltrated this data through email via SMTP protocol.

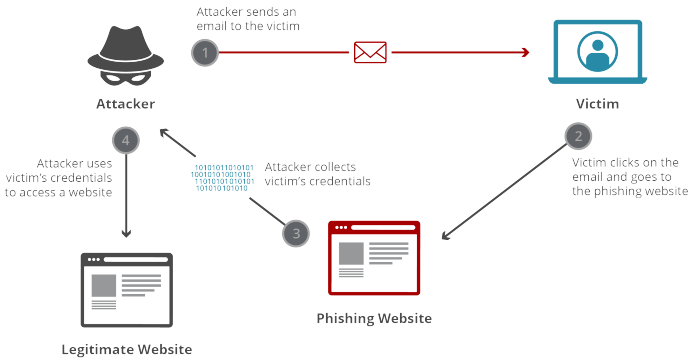


Figure 2. Process of Phishing Attack

Snake Keylogger

- The "Snake" Keylogger was identified in November 2020 and was noted to have the capability to steal credentials by capturing keystrokes, screenshots, and accessing the clipboard.
- The Keylogger uses the .NET framework and is spread through phishing campaigns where emails are sent containing office documents or PDF files. These attached files contain a macro that executes a Powershell command to download and install the keylogger.
- The keylogger sets itself to auto-run after initial installing so that no further action is required to execute the malware. After opening the document, the user is unaware that they have been compromised and may not realise for months, allowing the keylogger to collect significant amounts of data.

Agent Tesla Remote Access Trojan

- "Agent Tesla" was first discovered in 2014 and still used as recent as 2020. This malware was originally available on a Turkish website advertised as a remote access tool for personal use.
- Agent Tesla is a remote access trojan that specialises in stealing sensitive data, it is spread through phishing emails, hidden within attached documents. it uses the .NET framework and searches the system for web browsers which is can extract log in details from.
- Using SMTP to exfiltrate data through email, or even using telegram, the threat actor is able to gain access to the capture credentials.
- Due to the constant updates and developments of this malware including anti-VM and ant-sandbox capablites, it has managed to stay relevant by bypassing modern anti-virus and scanners for almost 10 years.

Conclusions

Throughout this project, I have explored the world of keyloggers, covering their functionality, their role in credential theft, and the potential security risk that they present. To mitigate these risks, individuals and organisations alike, should take the following measures to prevent falling victim to such attacks:

- Use secure password managers
- Regularly updating software including anti-virus
- Regularly updating knowledge of phishing scams

Future Work

If additional time was given, an in-depth analysis of real-world Keylogger malware samples would be carried out, this analysis would include:

- generation, delivery, and functionality
- static and dynamic analysis
- identifying unique features and behaviours
- determine most effective countermeasures for the specific sample

This would provide valuable insights for developing strategies to help prevent, detect and mitigate the sample in the future.

Acknowledgements

The Author would like to acknowledge the support of Dr. George O'Mahony and Dr. Hamdan Awan during the research and implementation phases of this project.

Smart Queueing System

Neil O'Sullivan,
BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

- QR code generation for individual queues.
- Queue management in a mobile app.
- Customers can join the queue by scanning a QR code.
- Customers receive queue information including position, estimated wait time, and estimated travel time.
- Location tracking to monitor a user's travel time from the appointment.

What does it offer?

Admin

- Mobile application.
- Generate a QR code for a queue.
- Manage all queues separately.
- Manage all users in the queue.

Customer

- Scan the QR code and join a queue.
- Receive queue updates.
- Share location and be notified when to walk back.

Things to note

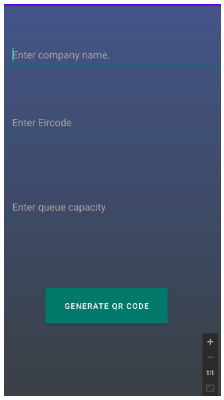
Low battery consumption

Real-time updates.

QR downloaded to the device.

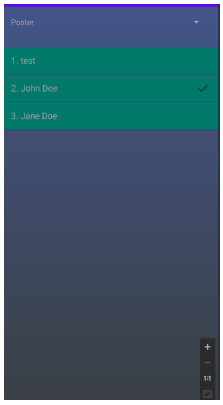
Application in action

Admin

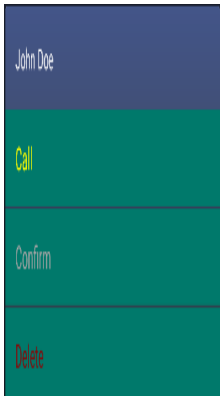


Generate QR Code

QR Code

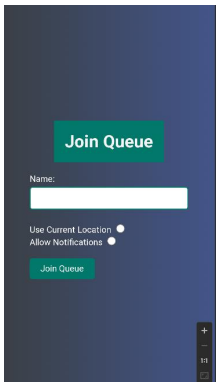


Manage Queues

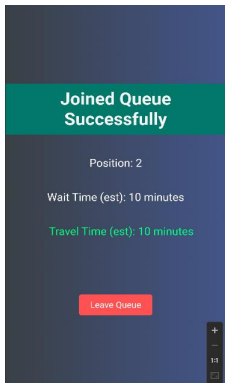


User Actions

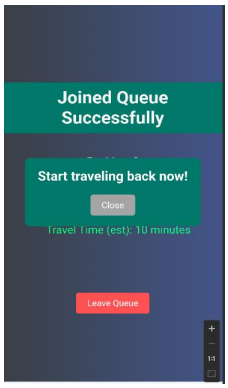
User



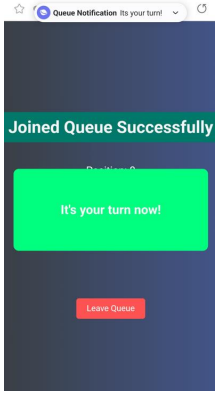
Join Queue



In Queue



Monitoring Location



User called

Technology

Android Studio · Firebase · JS · {RESTful API} · spring boot

Acknowledgments I would like to thank my supervisor for both semesters Dr. Alex Vakaloudis for his help throughout the year.

Automated Nutrition Assessment Simulator

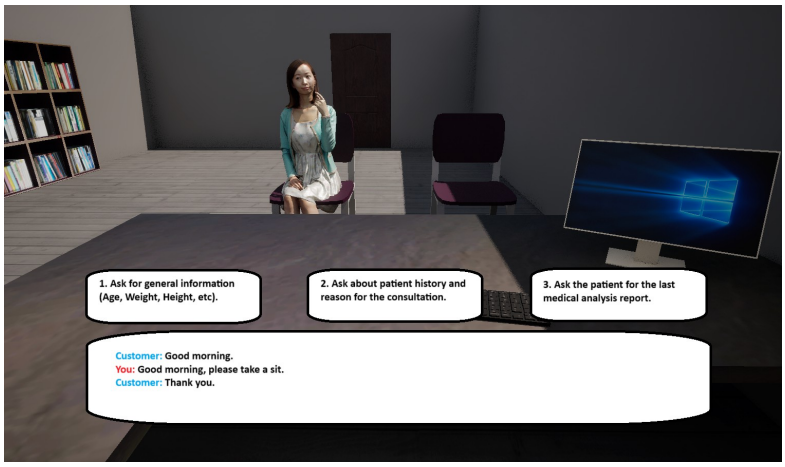
Jose Salas, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

This game serves as a simulation of a nutrition office consultation. This project explores the learning capabilities of serious games by providing the user a realistic environment in which he can put to practice their knowledge. The other purpose of the game is to create a middle ground between academic tests and reports and a real life problem.

Data Collection

The first part of the game consists of a serie of questions to gather the necessary information from the patient in order to assess what the problem is or how we can help them. This steps helps to put into the mind of the user the correct order and words to use when approaching a customer in a proffessional environment.



Input of Answers & Assessment

Acquired Information	Information Analysis
Name: Maria Leitz	BMI: _____
Age: 22	Energy requirements: _____
Height: 166 cm	Fluid requirements: _____
Weight: 39kg	Protein requirements: _____
Temperature: 35,6 ºC Heart rate: 68 BPM Respiratory rate: 14 BPM Blood pressure: 90/62 mmHg Tidal volume: 550ml Sodium: 142 mEq/L Potassium: 2.5 mEg/L Calcium: 8.2 mg/dL Phospate: 4.2 mg/dL	
History: Menses started at 12 but ceased 4 months ago.	

- All the gathered information is displayed in an easy-to-read manner.
- The user then enters the values that he/she has calculated for the different fields. The game finishes once all the fields have been entered.
- The game then compares the results provided by the user with the results of the calculations done in the background. The closer to the correct number, the more marks it has.
- 10% of the marks are reserved for making the questions in the correct order to ensure a humane treatment of the customer.

Conclusions

The main target of this project was to prove that virtually any type of paper examination can be digitalized. This results not only in an easier and less tedious way of learning but also in the better time management for the educational staff.

Acknowledgments

I would like to thank Alison O'Shea and Olivia Brickley for their guidance during this project. Also, Aoife McCarthy from the Nutrition Department in MTU for providing the necessary data.

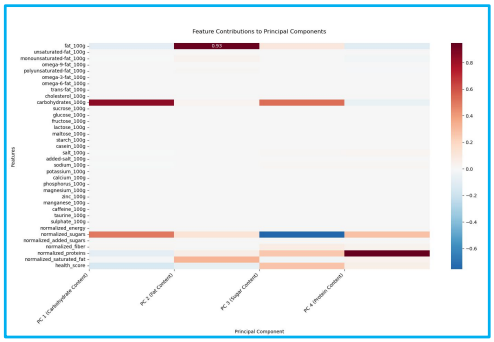
A Machine Learning Approach to Nutritional Analysis

Michal Strzelecki, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

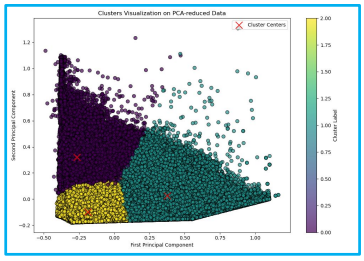
Introduction

To meet the growing need for educated dietary decisions in the era of convenience, the project uses mobile and machine learning technologies to provide immediate nutritional insights. The goal is straightforward, to provide consumers with an intuitive Android app that uses barcode scanning to group food items into clusters that are health-conscious. Through the conversion of complicated nutritional information into easily comprehensible health ratings, encouraging consumers to regularly integrate healthier food choices into their everyday purchasing experiences.

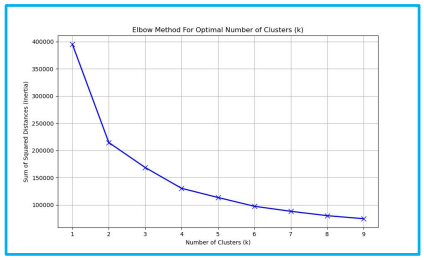
Clustering



Heatmap represents the contribution of each feature to the principal components



Clusters Visualization on PCA
(Principal Component Analysis)

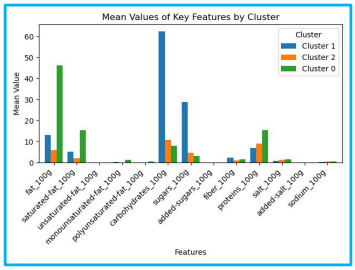


Elbow method explaining optimal number of clusters.

The clustering results supply the recommendation engine within the app, which helps users receive basic nutritional advice based on the product scanned. By defining these clusters, dietary decisions can be addressed in a more refined way, with data serving as the foundation for health optimization.

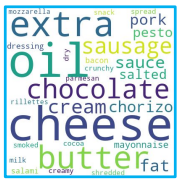
Sample Analytics

The bar chart showcases the model's feature analysis, with mean values to highlight the significant nutritional differences while the word cloud for individual clusters helps us to identify most common words from product name. Combining everything together we are able to more informally make a decision regards which Cluster belongs in which food type category (Good/Bad, consume in Moderation)

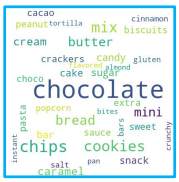


Sample graph of summary statistics

Generated Word Cloud from results of NLP (Natural Language Processing) from individual Clusters



Cluster 0 (Moderate Consumption)



Cluster 1 (Unhealthy)

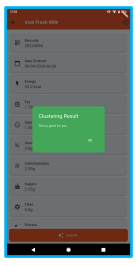
Cluster 2 (Healthy)

Application

Main Menu

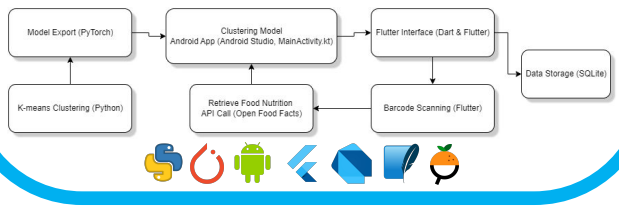
Manually adding product

Product History



Detailed Product History

High-level architecture diagram



Conclusions

Integrating machine learning models for data clustering into mobile applications can greatly simplify nutritional information, making healthy choices accessible and straightforward. The successful development and implementation of the barcode scanning feature, coupled with the nutritional clustering model, underscores the potential of technology to positively influence public health. By providing a clear classification of food items into health-oriented clusters.

Acknowledgments

Special thanks my wife and child for their patience.

Brain Tumor Classification using ResNet50v2-GRU Deep Learning model

Muhammad Umar Tariq,
BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

Introduction

Brain tumors remain a challenge in healthcare, posing a threat to the lives and well-being of those affected. Timely and accurate diagnosis is of great importance, as it directly impacts treatment options and patient outcomes. The existing diagnostic methods, while effective, require extensive manual effort and can be subject to human error. As such, there is an urgent need for innovative, technology-driven solutions to enhance the capabilities of healthcare professionals in detecting and classifying brain tumors.

Dataset

The models are trained and evaluated using the publicly available **'Brain Tumor MRI Images'** dataset.

Training

- Glioma – **1321** images
- Meningioma – **1339** images
- Pituitary – **1457** images
- No Tumor – **1595** images

Testing

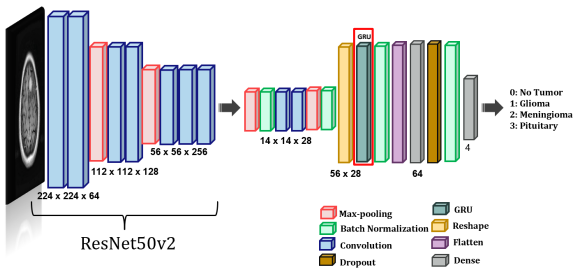
- Glioma – **300** images
- Meningioma – **306** images
- Pituitary – **300** images
- No Tumor – **405** images

Research Questions

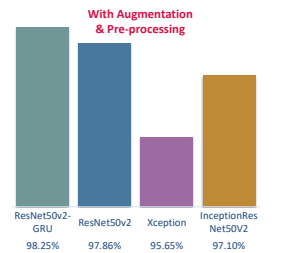
- How does the proposed hybrid model perform compared to transfer learning model for tumour classification?
- How does data augmentation and preprocessing effect the accuracy of classification models?
- How does the model perform on the proposed dataset?

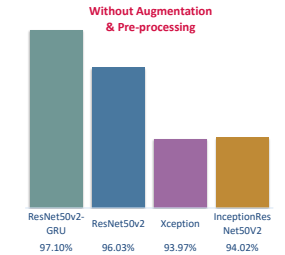
ResNet50v2-GRU

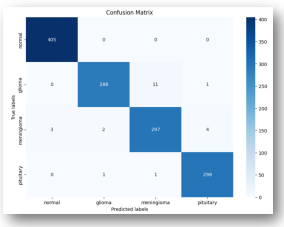
This project introduces a novel approach to brain tumor classification utilizing a hybrid **ResNet50v2-GRU** model. ResNet50v2, a transfer-learning CNN architecture along with 2 additional convolution layers, is used to extract features from images. The pre-trained ResNet50v2 layers are frozen and the additional layers added to the architecture are fine-tuned. The GRU layer is introduced to discern patterns within the extracted features. This adds the novelty to the architecture and enhances its classification capabilities.

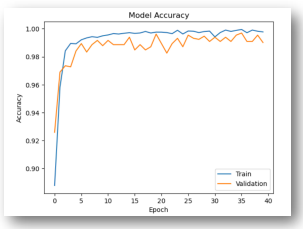


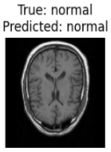
Benchmark Models & Results

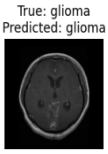


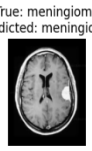












Data Augmentation & Pre-processing

Data augmentation and image pre-processing was applied to all the samples in training set. This was done to expand the training set and add variety to it.

- Brightness Adjustment
- Contrast Adjustment
- Rotation of 20 degrees clockwise
- Rotation of 20 degrees anti-clockwise
- Salt and Pepper Noise

Conclusions

The proposed model outperforms other Deep Learning models used in this domain.
Adding data augmentation and pre-processing increased the accuracy from 97.10% to 98.25%
This shows the potential of hybrid models in enhancing diagnostic accuracy and advancing medical image analysis.

Acknowledgments:

The author would like to thank Debora Pereira Salgado, Christian Beder and Brian Murphy for their supervision, and Stryker for sponsoring this project

Matching Food Brands with Food Brand Database

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University
Thando Tsebedu, BSc Honours in Software Development
Department of Computer Science,
MTU Cork, May 2024

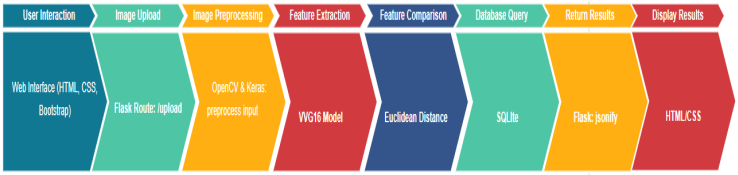
Research Question and Objective

The investigation seeks to understand if automated image matching can streamline dietary data collection. The goals include creating an accurate algorithm, ensuring user-friendly interactions, and accommodating various image conditions.

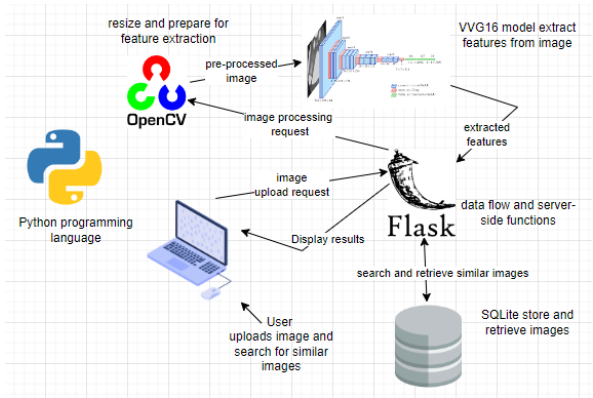
The aim is to develop a highly accurate image matching system that simplifies the process of data collection of food brand images. We focus on delivering a user-friendly platform that contributes to better brand recognition and aids nutritional research.

Methodology

1.User Interaction: Easy web access for image matching.
2.Image Upload: Secure Flask-based image storage.
3.Image Preprocessing: Consistent pre-analysis image setup.
4.Feature Extraction: VGG16 model extracts key image details.
5.Feature Comparison: Euclidean distance gauges image likeness.
6.Database Query: Quick SQLite search for matching features.
7.Return Results: Flask turns data into JSON for the frontend.
8.Display Results: Matched images displayed via HTML/CSS.



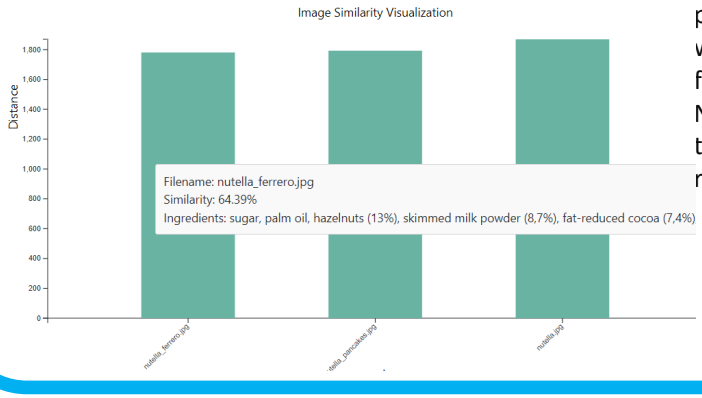
Systems Architecture



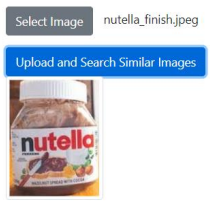
VVG16 Model

Using the VGG16 neural network, our application extracts complex features from images to create unique signatures. These signatures are then compared to find visually similar images within our database. Flask manages the flow, allowing users to upload images and receive instant results.

Search Results



Histogram shows the similarity of each product to the (Nutella) uploaded image, with lower bars indicating closer matches. It found matches ranging from very similar Nutella jars to less similar items, demonstrate the system's ability to identify a variety of related food products from our database.



Conclusion

This project not only bridges the gap between digital image recognition and food brand identification but also sets the stage for future advancements in the field. The success of the system highlights the potential for further research and development.

References

1. Blurredmachine, "VGGNet-16 Architecture: A Complete Guide," Kaggle, 2020. [Online]. Available: https://www.kaggle.com/code/blurredmachine/vggnet-16-architecture-a-complete-guide

Acknowledgments

Special thanks to Dr. Christian Beder, whose expert guidance significantly shaped the implementation of my project. I am also grateful to Dr. Alison O'Shea and Dr. Brian Murphy for the expert advice and support they have provided throughout this journey.