By Rizwan Ye
Supervisor: Shoujin Wang

Enhancing XAI Interpretability: Using Radar Charts to Simplify Feature Based Explanations and Reduce Cognitive Load

XAI

Acknowledgements

I would like to thank Shoujin Wang he is a professor at UTS for his guidance throughout my honours project and most importantly to be freely available to supervise me in a short notice, he has helped me tremendously to finalise my project and allowed this project to be accepted and provided me resources that I should review to make this project into a presentable report. Without his help this project would not be finalised.

I would also like to thank Jainlong Zhou, he is an assistant professor at UTS and has tremendously helped me throughout my degree in allowing me to understand the concepts of XAI and cognitive load, his guidance allowed me to discover XAI and cognitive load. With his guidance in the past two years, I was able to fine tune the original aims of my research to a final presentable paper at the end of my degree. Jianlong has provided me with relevant papers and novel ideas to pursue during my endeavour. He has made many suggestions helped clarify some complex concepts He is a long-time contributor in this space and coauthored many papers in this domain.

In terms of the formatting of this report, I would also like to thank James John Geoffrey want's report on his interpretability of machine learning models in traffic prediction. Apart of my structuring report design was inspired based on his project.

Abstract

Explainable Artificial Intelligence (XAI) is an emerging field that aims to develop machine learning models and algorithms that can provide transparent and interpretable explanations for their decisions and predictions. One of the main challenges in XAI is the effective communication of these explanations to non-experts while also minimizing cognitive load.

As society becomes more dependent on intelligent complex systems, the prevalence of black box models poses a significant real-world challenge, leading to issues in both social and environmental contexts. As a result, the development of Explainable Artificial Intelligence (XAI) models that offer holistic interpretability for both data scientists and end users become more increasingly important.

One of the primary issues facing the field of XAI is that most researchers concentrate on the technical implementations of XAI models; however, little research has been done on how these models are implemented in the real world and interacted with by the end user or how their explanations affect end users. Thus, showing that this area is often overlooked within the field of XAI.

To address the need for improved interpretation and understanding of XAI results, this study investigates the impact of different types of explanations from XAI models by measuring the cognitive load of end users. Specifically, we conducted an empirical survey to explore the effectiveness of text-based explanations, traditional graphical explanations from commonly used XAI models, and Radar charts.

The results revealed that Radar charts received more favourable ratings from the general audience compared traditional XAI explanations. However, text-based annotations proved to be more effective in certain situations. A noteworthy finding was the widespread difficulty participants experienced in comprehending traditional XAI visualisations without adequate context or knowledge in this domain.

Building upon these results, we developed a prototype loan approval web-based application that implements the findings of the empirical survey. The prototype aims to demonstrate the feasibility of using radar charts to visually explain feature-based outputs from XAI algorithms. Additionally, a language model was implemented as well to generate AI explanations on the outcomes to improve interpretability and explainability.

The findings of this paper provide insights into how the general public reacts to current XAI explanations, as well as new approaches for investigating appropriate visualisations in the realm of explainable AI. Furthermore, by implementing a XAI prototype based on the empirical survey results, it demonstrates the feasibility of XAI machines being used in the real-world environment, demonstrating their ability to improve ethical decision-making and foster a deeper understanding of AI/ML programs for non-technical users.

TABLE OF CONENTS

Table Of Figures

Table Of Tables

Nomenclature

LIME – Local Interpretable Model Agnostic Explanations

ML – Machine Learning

AI – Artificial Intelligence

XAI – Explainable Artificial Intelligence

1.INTRODUCTION

Globally, industries increasingly rely on intelligent systems for decision-making. However, over-reliance on these systems can lead to biased conclusions, especially when trained on flawed data. The banking industry, in particular, employs complex AI mechanics to determine loan eligibility, making it susceptible to such issues (Caron, 2019).

To address these complexities, research in Explainable Artificial Intelligence (XAI) seeks to enhance understanding of AI models due to their black-box nature (Ali et al., 2023). Improved transparency in these models fosters better interpretation of the decision-making process, promoting ethical fairness and explainability (Alaa, 2021).

From a holistic perspective, XAI outputs should cater to all stakeholders, optimising cognitive load for transparency and explainability (Langer et al., 2021). However, current XAI explanations are often understandable only to those with domain knowledge.

Furthermore, those with technical knowledge in this domain are also focused in improving the technical aspects of XAI algorithms rather than focusing on improving the explainability of these machines for end users. As a result, emerging research, such as visual analytics (VA), can potentially improve this gap within XAI research, as it explores how enhanced visualisations can improve comprehension for end-users (Alicioglu & Sun, 2021).

To determine the appropriateness of a visualization, it must produce accurate results, promote learning, and maintain an optimal cognitive load (Padilla et al., 2018). Understanding the positive and negative impact of the visualisation requires research into its cognitive load and usability (Huang et al., 2009).

This study aims to assess the comprehension of traditional eXplainable Artificial Intelligence (XAI) explanations and determine whether radar charts can provide a better visual explanation through the use of an empirical survey. Radar charts were chosen in this context as they are commonly used in finance and engineering to visualise feature-based scores. However, radar charts are not commonly discussed or utilised within the field of computer science.

Furthermore, to implement the results of the empirical survey, a web-based prototype loan application that utilizes an XAI algorithm is developed to produce radar chart visualizations. Additionally, a language model is integrated to generate automatic explanations. This development aims to demonstrate the capability of the system in a real-world environment.

This report will begin with a literature review to explain the current dynamics of the XAI domain, then delve into the project aim and scope, methodology of the project which discusses stake holders and requirements, survey development, and web application implementation based on the survey results. To conclude the report, at the end the discussion and summary will draw together the key findings and insights.

2. Literature Review

In order to understand the context of this project, understanding the 3 main concepts listed below is important and interrelated within the context of XAI programs and how it affects an end user's interpretation of these designs. In this section it reviews prior research paper studies in the field of cognitive load and AI and how they align to this project.

2.1

Explainable Artificial Intelligence and Visual Analytics
There are two main areas of enhancing transparency of AI models: improving the simplicity of AI models and providing post hoc explanations for AI decisions. However, there is a lack of research involved with improving post hoc explanations, namely enhancing the visual component of post hoc explanations which is crucial to improving user understanding.

2.2

Radar charts As a Visual Tool to Represent Data
Radar charts are an under-researched but potentially beneficial post-hoc visualisation that could be used. This type of visualisation is commonly used to represent multidimensional data in the business, gaming, and financial industries. The familiarity of radar charts, as well as their ability to represent feature-based data, may make them a potentially useful visual tool within XAI post-hoc visualisations.

2.3

Cognitive Load and XAI Interpretability
Data scientists in the XAI field tend to prioritie model performance over user understanding, resulting in ambiguous explanations. Cognitive load, a measure of mental effort required to comprehend information, has a significant impact on system comprehension. High cognitive load can impair model comprehension and perceived feasibility. Thus, providing transparent explanations for XAI models with optimal cognitive load is critical for building trust and facilitating better model interpretation decision-making.

2.1 Explainable Artificial Intelligence and Visual Analytics

Explainable artificial intelligence (XAI) is a rapidly growing field, especially in the context of ensuring ethical AI practices. The primary aim of XAI is to enhance transparency and interpretability within AI model decision-making processes (Ali et al., 2023).

Within the domain of Explainable Artificial Intelligence (XAI), two main areas of focus emerge which is the enhancement of inherent simplicity in AI models and the implementation of post hoc explanations (Vilone & Longo, 2021). Post hoc explanations involve generating justifications for AI model decisions after the model has made its predictions. It's crucial to note that post hoc methods don't modify the model itself but rather retrospectively analyse its behaviour.

A significant portion of XAI research is dedicated to the development of post hoc solutions. This emphasis is driven by the increasing need to incorporate interpretability into existing AI models as many industries have already deployed AI systems constructed upon their internal intricate architectures, such as deep learning, which were not originally designed with interpretability in mind (Linardatos et al., 2020). Thus making it difficult for these industries to modify their model intrinsically.

In light of these challenges, post hoc methods become an attractive solution, particularly in industries like banking, where modifying the inherent model of their system can be a costly endeavour (Ali et al., 2023). By opting for post hoc approaches, organizations can achieve interpretability without directly impacting the underlying model, making them a practical choice in scenarios where system modification is economically or technically challenging.

The methods proposed in XAI post hoc models are primarily tailored for data scientists. However, from an end user's perspective, understanding the decisions made by post hoc models can be challenging and significantly impact their perception of the model (Ali et al., 2023). Since the end user primarily interacts with the post hoc explanation, a convoluted layout combined with technical visual components can negatively affect their perception of the model (Linardatos et al., 2020). Therefore, improving the visual analytics component of a post hoc explanation is essential to enhance the interpretation and meaning of the output.

According to Shankar & Duraisamy (2018), from a visual analytics standpoint, optimising a visual explanation requires an appropriate representation of the data that should cater to all stakeholders as visual analytics is highly interdisciplinary. Assuming that post hoc explanations are primarily feature-based, the focus should be on identifying and displaying these feature components through a graph that is tailored towards feature explanations and is familiar to understand.

2.2 Radar Charts as A Visual Tool to Represent Data

Various visualizations can be employed for post hoc explanations; however, not all may be ideal or suitable for the intended target audience. One potential visualization that could align with the cognitive load of the end users is the radar chart, also known as skill graphs. This style of visualization is a prevalent technique utilized across fields including engineering, gaming, and many more.

One of the main reasons why radar charts are widely adopted across these industries is their ability to represent multiple dimensions or variables in a single, easily interpretable format, especially due to their lower visual search time compared to traditional line and bar graphs (Abeynayake et al., 2023).

In the context of XAI, where complex machine learning models generate outcomes that may not be immediately understandable to non-experts, visualising both positive and negative feature importance values of post hoc outputs appropriately is important (Alicioglu & Sun, 2021).

Radar charts can offer the potential to visualise these feature importance scores suitably in this context. However, radar charts or innovative visualisations are not commonly discussed in the field of XAI.

According to Seide, Jensen, and Kieser (2021), they have similarly conducted research on radar charts' feasibility within the domain of visualizing simulation and estimation results in network meta-analysis, based on their findings they have demonstrated that radar charts hold characteristics that can represent data that contains complex high-dimensional data structures. This study demonstrates how radar charts are steadily being researched as potential visualisation to justify complex multi-dimensional data.

Similarly, Morales-Silva et al.'s (2020) research on the use of radar charts as a visual tool within the healthcare domain demonstrates the capability of radar charts in effectively visualising sensitive data and presenting it to a diverse audience. Most importantly, their research suggests that radar charts allowed the unique juxtaposition of data.

Extending this understanding to the field of Explainable Artificial Intelligence (XAI), it becomes evident that the use of radar charts could offer a valuable visual framework for presenting complex data outputs in a comprehensible manner for the majority of stakeholders utilising ML programs, which can be beneficial for non-technical end users.

Thus, by drawing on other disciplines the insights and implementation of radar charts demonstrated both in Seide, Jensen, and Kieser (2021) and Morales-Silva et al.'s (2020) paper suggest that radar charts may be a suitable tool for visualising featured-based post hoc data within the realm of XAI.

2.3 Cognitive Load and XAI Interpretability

In the field of Explainable Artificial Intelligence (XAI), data scientists often prioritize enhancing AI model performance, sometimes neglecting the crucial perspective of end-users. This oversight results in ambiguous explanations provided by XAI post hoc visualizations, as noted by Saeed and Omlin (2023).

A recognized measure of a system's perceived understandability is cognitive load (De Jong, 2010), commonly employed in psychology to quantify the mental effort required for individuals to process information. According to De Jong (2010), cognitive load comprises intrinsic load, inherent to material complexity, and extraneous load, related to material presentation and organization.

Numerous studies such as hao Xiaolin et al, (2023) have suggested that increased extraneous cognitive load negatively impacts individuals' understanding and learning experience. Thus, a well-structured model explanation that has low extraneous cognitive load is crucial for enhancing XAI interpretability for end users, which can help build trust and contribute to improved decision-making in understanding the model (Ali et al., 2023).

According to Ouwehand et al, (2021), cognitive load is often evaluated using the Likert scale method. In this approach, participants, engage in problem-solving tasks and provide nuanced ratings on a scale of 1 to 10. This scale captures the intricate interplay between their perceived mental effort and the difficulty of the task at hand. The Likert scale, with its numerical representation, aims to gauge how users introspectively assess the cognitive demands imposed by a learning task.

Among the current literature in this field, measuring cognitive load for XAI explanations was not commonly researched. However, in the field of visual analytics, there is a great emphasis on cognitive load and visual analytics research.

One noteworthy study by Huang et al. (2009) utilized this methodology to explore cognitive load in users' understanding of various graph visualizations. Based on their research, they investigated factors influencing cognitive load, comparing its impact across different visualization types. The study demonstrated the importance of minimizing visual cognitive load while maintaining reasonable overall cognitive load for enhanced information retrieval and processing.

As a result of their research, they discovered that cognitive load can be used to determine the appropriateness of a graph in a given context. However, they emphasized the importance of generalizing results. Building on their research and applying it to the XAI domain, their paper suggests that using the Likert scale and cognitive load theory to measure the effectiveness of a graphical representation can help establish a link between user understanding and their interpretation of XAI visualizations by.

3. Project Aim and Scope

This project was conducted in 2023 as part of the University of Technology Sydney honours program. Before the commencement of this honours project, I had worked closely with Jianlong Zhou, who specialized in the domain of XAI. The main goal was to research novel XAI visualizations to help identify possible visualisations that can help improve the cognitive load of XAI outputs.

During the initial stages of the analysis, I identified informational gaps and a lack of research involving the end-user side of these XAI programs, particularly the visually explainable side. Multiple visualizations were tested and developed before this final project, and the most promising one was a radar chart. As a result, the final research focused on utilizing radar charts to improve the understandability of post-hoc explanations for end-users. To assess its efficacy, an empirical study needs to be conducted. Furthermore, the development of a prototype platform will demonstrate its application in a real-world environment.

As the majority of banks in the real-world environment deploy complex models to determine the loan approval process, the context of this project is situated around the financial domain. Consequently, this study was conducted to demonstrate how the banking industry can improve its transparency within its domain.

3.1 Aims and Objectives.

The thesis is as follows:

> That improving the visualisation of XAI output programs improves the end user experience and compression when dealing with complex AI programs which improves transparency and understanding.

The main goal of this project is to be application-specific to demonstrate a new visualisation tool for XAI outputs, where we will investigate the current cognitive load experienced by end users in traditional explanations used within the XAI domain and based on the empirical study results, develop a prototype web application implementing these solutions. This will be addressed in stages as follows:

1. The first step is to deploy the necessary survey of the explanations in both online and offline mediums, allowing end users to interact with the survey to score their cognitive load.

2. The second phase involves developing a prototype website that implements the results of the survey. This step aims to demonstrate the empirical results within a real-world environment.

3.2 Scope of the project

When considering the project's scope, my initial plan was centred around studying potential visualisations and assessing their cognitive load for end users. However, I realised that this approach lacks practical applicability.

As a result, for the final project, I intend to develop a web-based eXplainable Artificial Intelligence (XAI) platform which was inspired by IBM’s AI Explainability 360 design. This platform will be interactive for end users and will visualise feature-based outputs of the LIME XAI output using a tailor made radar chart and automatically provide explanations of the result.

3.3 Project limitations

Since I will be launching an empirical study and developing a web-based prototype, the project has a large scope. As a result, time and knowledge are limited.

Furthermore, because XAI is a new domain, there is limited information and resources to rely on, and covering all domains of XAI and the types of modules it offers will be difficult. As a result, this project will only use the well-known LIME XAI package.

Another limitation is that prioritizing the accuracy of the backend machine learning was not feasible given the project's tight deadline and limited timeframe. As a result, the final classification accuracy may be low. This trade-off was deemed acceptable because, rather than prioritising classifier accuracy, the extra time was able to be utilised in building a functional prototype that implements an XAI module with a radar chart application and a language model explainer. Thus, emphasising the application-specific objective of this project.

In terms of the language model used, this project utilised the Orca Mini from GPT4ALL to generate explanations, rather than ChatGPT API due to budget constraints. As a result, the language model speed and depth rely on the localised machine’s performance. Thus, limiting its explanation capacity.

Lastly, in terms of the empirical survey, gathering survey participants was difficult as a result the survey participant enrolment may not be as high or standardised as originally expected.

4. Project Methodology

This section will describe the dataset that will be used, the stakeholder analysis involved in the project, the project requirements, and the technology that will be used to conduct the empirical survey and develop the web prototype.

4.1 Stake Holder Analysis

The project's main goal is to provide insights into how current experimental XAI programs can be implemented to help end users better understand the outcomes of complex AI/ML programs. To comprehend how the user interface should be deployed and what it should look like, and what type of visualization should be used, we must first identify the end user's requirements.

Based on the IBM Explainability 360 design (Arya et al., 2022), the researchers proposed that the architectural design of an XAI model should be flexible and easy to use, catering to both data scientists and end users with no technical knowledge. As a result, to build a visualisation and appropriate web application we will need to use a design that is familiar and easy to use.

Looking at the banking industry we can summarise that the end users utilising these applications are primarily the customers, followed by call centre operatives which are managed both by the bank owners and regulators.

As a result, based on this finding, we can assume that the end users have the following requirements:

- Customer: Normally, customers who apply for loans or credit cards need to input their information into the application. This process would be entirely automated by a backend ML machine to determine whether they meet the criteria. Thus, end users may want a generated visualization that is easily interpretable without clutter and an in-depth explanation of the results.

- Call centre: When end users want to talk to a bank, they would normally call a call centre employee to identify and clarify their application. Normally, they want to have a clarification of why their application went through or why it was declined. As a result, employees may also want a quick overview of the visualization and explanation due to their call time being tracked. Thus, call centre employees may also want an XAI output that is easily and quickly understood.

- Bank Managers and Owners: The upper management of these organizations tends to want a more holistic approach when working with complex algorithms to avoid biases and incorrect classifications of approvals. However, as they are not directly involved in the development, the bank managers and owners may want a design that fits the needs of the consumer to help with transparency.

- Regulators: External regulators of Australian banks normally oversee the process of banks to determine if they are following the right process. For example, they want the reasoning of why the bank has approved or declined an application. As a result, regulators may want a visualization and explanation of the outcome that is holistic and easily understandable for all stakeholders to avoid discrimination.

As a result, to appropriately maintain the stakeholder requirements and consider the needs when implementing the design of the web application and visualisation, we must maintain the stakeholder requirements as listed below.

Stakeholder	Interest	Influence	System Needs & Requirements
Customer	High	Low	Requires transparent and explainable outcomes of the system. End user requires a low technical and cognitive load explanation of their outcome.
Call Centre Operator	Medium	Medium	Call centre operators are not technical employees; thus, they require an output that is easily understandable and quickly generated to reduce their call time.
Bank Owners and Managers	High	High	Bank owners and managers focus on KPIs such as application approval rate and employee satisfaction. Thus, they want to reduce customer pain points and help satisfy end user's needs.
Regulators	Medium	High	Regulators of banks focus on ensuring compliance and transparency in the financial system. They require comprehensive and detailed explanations to evaluate the system's adherence to regulations.

Table 1 Stake Holder Analysis Table

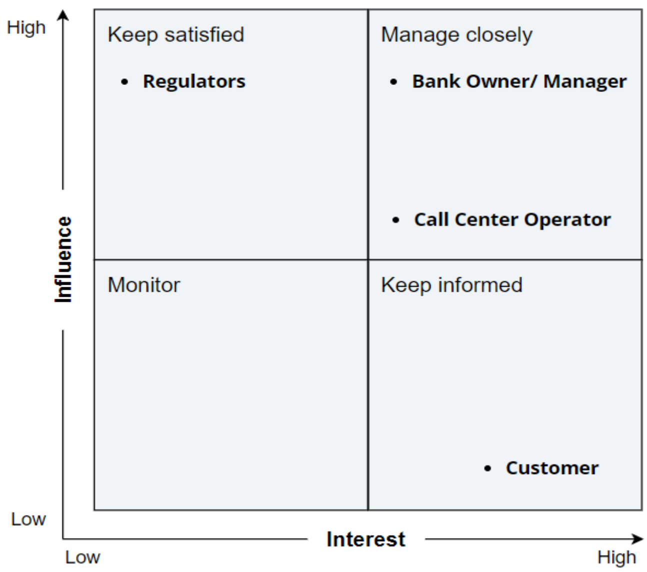


Figure 1 Stake Holder Analysis Grid

4.2 Requirements Gathering

Having an accurate and reliable requirement is pivotal for any organization developing new software for its systems. In the diagram below, it displays a stakeholder requirement grid. These types of grids help us understand how the components react when we utilize them.

To fully understand the requirements of this program, Figure 2 demonstrates how the system will work.

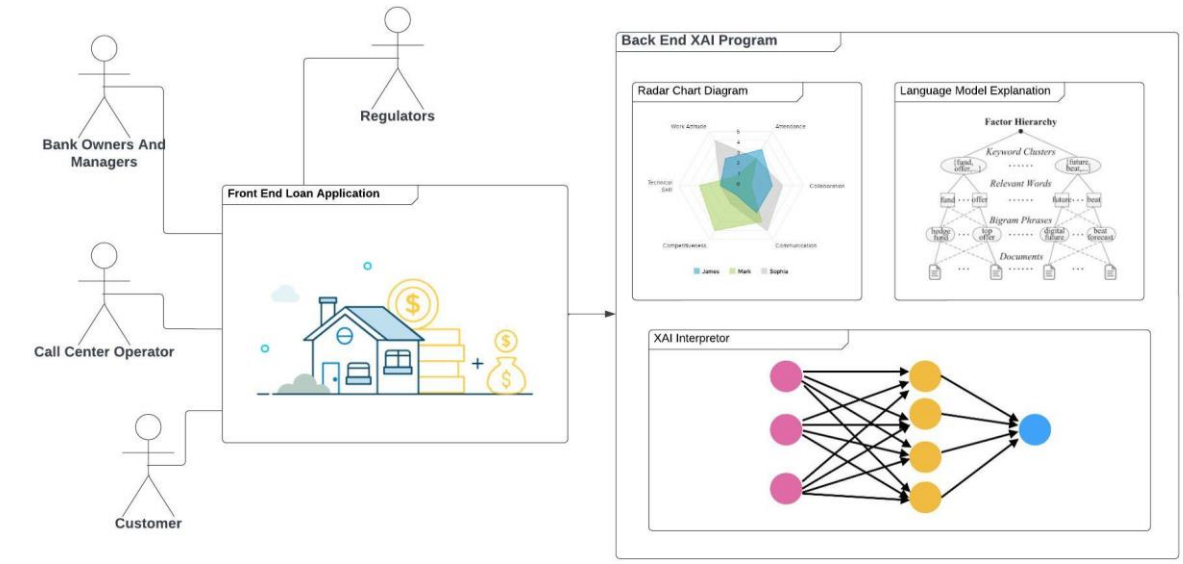


Figure 2 Model Interpretation Boundary Diagram

Component	Requirement
Actors (End users)	End users should be able to interact directly with the Front-End Loan Application platform.
Front End Application	The front-end application should react with the user input and interact with the back-end XAI Program to provide explanations to the end user
Back End XAI Program	The back-end XAI Program should have an XAI interpretation built within it to output explanations and a tailored description to the end user
	The XAI visualisation should utilise a low cognitive load design for the end user
	The XAI program should use a combination of visual-based and text-based explanations for the end user

Table 2 Stake Holder Requirements

4.3 Technology Utilised

To understand how the project was developed, below lists the type of technology used in this development.

Hardware Both the XAI program and machine learning model were developed and trained using my Alienware m15 r5 which utilised Windows 11 that had a core i7 with 16 GB RAM.	
Python Python programming is a popular programming language used in data science; this was primarily used to develop the model. Namely, packages such as Sklearn, Lime, and Langchain were used.	
Streamlit To launch the web application Streamlit was utilised. Streamlit is a free and open-source application that allows the deployment of data science projects.	Streamlit
GPT4ALL GPT4ALL is an open-source platform that provides a collection of language models trained on a massive corpus of data, allowing localized running of these language models on your computer. This was utilized for the automatic explanations of the XAI program outputs.	GPT 4All
SurveyMonkey And Google Forms To encourage user interaction with the survey for the empirical study the survey was launched on an online platform called SurveyMonkey and also using Google Docs.	SurveyMonkey
GitHub GitHub is an online coding platform that allows you to utilize and share code online. It was utilized to back up the code iteratively as I progressed in the programming stage.	GitHub

Table 3 Technology Stack Table

5. Survey and Visualisation Development

In this section, I will cover how the visualization and interpretation were developed and how the empirical survey was launched.

Based on the introduction of the report, the current environment faces a problem in creating an interpretable and understandable XAI (Explainable Artificial Intelligence) or VXAI (Visual Explainable Artificial Intelligence) for end users without a technical background. One of the limitations is the lack of easily understandable designs. As a result, there is a need for interpretable designs that suit the intended target audience.

To evaluate the best interpretations, we must consider a holistic approach when implementing appropriate design methodologies used in both visual analytics and psychology. To do this, we need to rely on the leveraging principles of visual perceptions, such as Gestalt principles, which can help organize and structure information in visualizations to promote comprehension and highlight key patterns or relationships (Yalcinkaya & Singh, 2019).

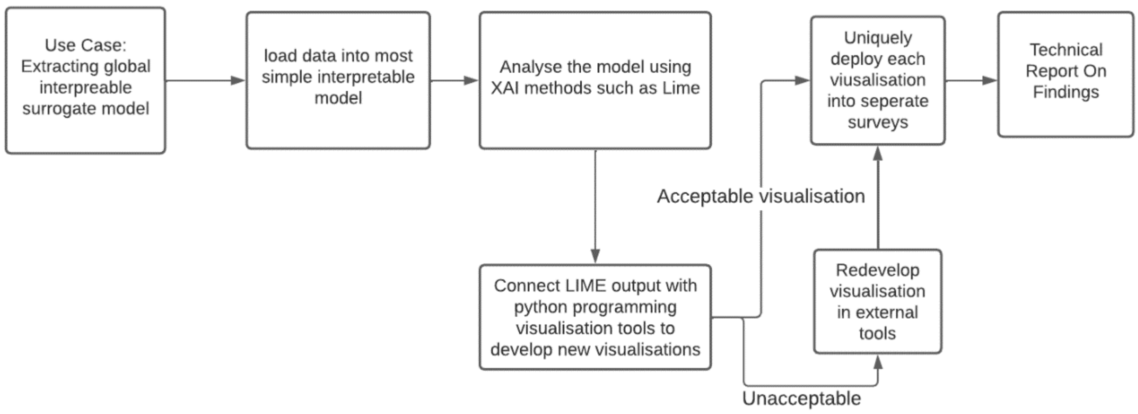


Figure 3 Flow Chart Design of Empirical Survey

Figure 3 visualizes the journey of conceptualizing the radar chart visualization. This diagram serves as the foundation for the original design methodology in identifying the appropriate visualization and outlining the design formation process. Before this honours research, several visualizations were developed, namely parallel coordinates and scatter graphs, to depict feature-based data points. However, despite the innovativeness of these visualizations, they were perceived to impose a high cognitive load based on our interpretation of the design.

As a result, in this stage of the project, we are finalizing the technical report findings from the previously launched survey. The survey launched aimed to test the cognitive load associated with individuals' interpretation of annotation explanations, feature-based graphs from the Lime XAI model, and the radar chart diagram.

5.1 Data Exploration and Pre-processing

In terms of the dataset utilized, this project used the "Loan Prediction" dataset from Analytics Vidhya. However, this dataset was limited to only 615 data points, thereby restricting its data capacity.

Given the time constraint involved with this project and that accuracy was not a high priority, we opted to utilize this dataset as the primary source for training the model and to be utilized as part of the back-end design of the XAI web application.

In terms of pre-processing, the Loan_ID and Credit History were dropped from the dataset. Additionally, the data types containing categorical data were converted to nominal or ordinal data types to prepare for the machine learning algorithm, using one-hot encoding. As the main goal was not an accuracy rate, this initial step was not given full focus.

Columns	Description
Loan_ID	A uniques loan ID
Gender	Male/Female
Married	Married(Yes)/ Not married(No)
Dependents	Number of persons depending on the client
Education	Applicant Education (Graduate /Undergraduate)
Self_Employed	Self emplyored (Yes/No)
ApplicantIncome	Applicant income
Coapplicant income	Coapplicant Income
LoanAmount	Loan amount in thousands
Loan_Amount_Term	Term of lean in months
Credit_Hostory	Credit history meets guidelines
Property_Area	Urban/Semi and Rural
Loan_Status	Loan approved (Y/N)

Table 4 Representation of Dataset

5.2 Machine Learning Algorithm Development

To utilize the XAI algorithm, a machine learning algorithm must be implemented as the baseline model before an XAI algorithm can be implemented. For algorithm development, a simple decision tree was selected rather than a deep learning model. This choice was influenced by research papers such as Zhang et al. (2022), which emphasized the importance of maintaining transparency in the entire platform to create an interpretable surrogate model.

Following this rationale, the decision tree was used in this case to conduct a simple binary classification task in determining whether an individual was eligible for a loan, with the Loan status chosen as the target variable which was binarized as 1 for "yes" and "0" for "no," while other attributes were converted to feature variables.

The decision tree was then trained on a 70/30 split using the DecisionTreeClassifier function within Python, achieving an accuracy rate on average of between 65% - 71.89%. Given that accuracy and validation were not part of the scope, this was deemed reasonable for this study.

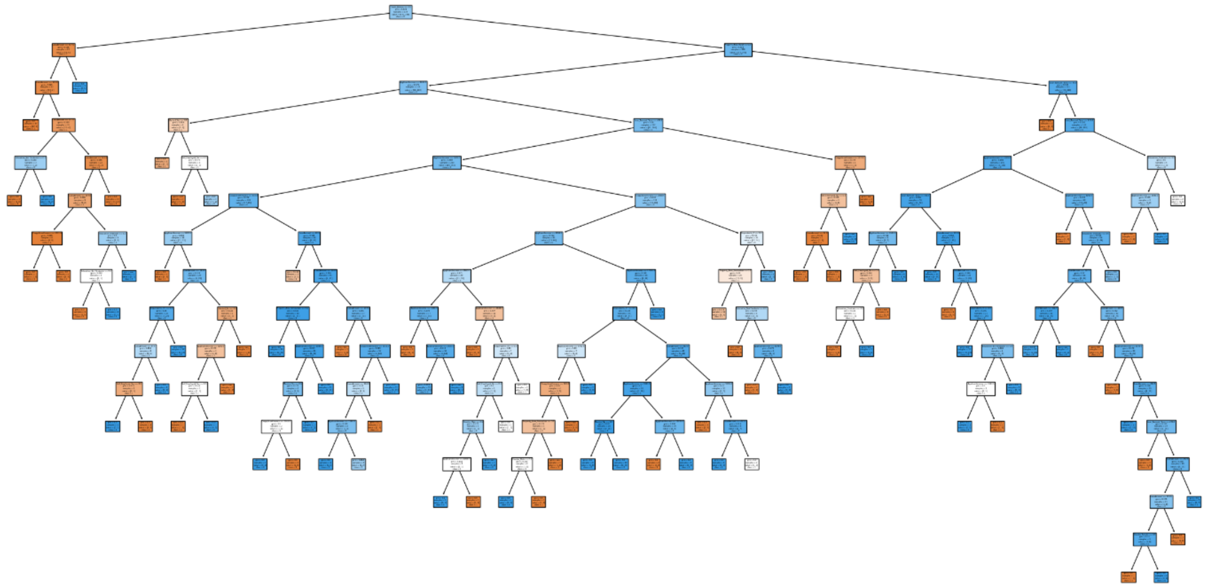


Figure 4 Decision Tree Visualisation

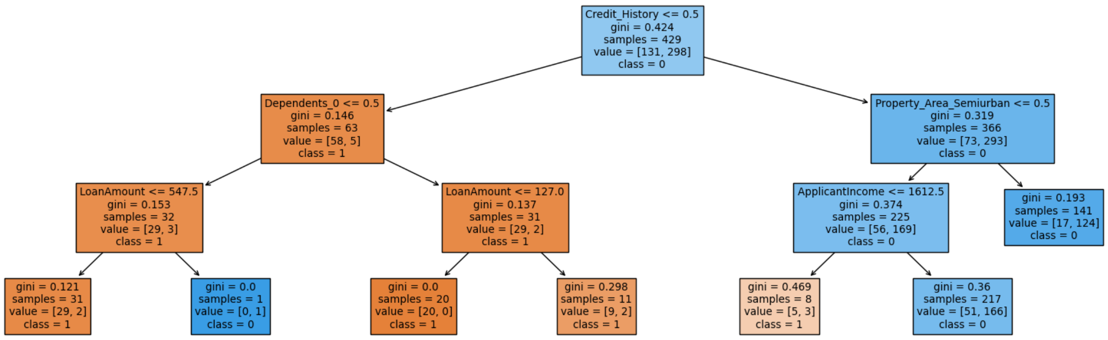


Figure 5 Pruned Decision Tree Visualisation

The visualisation above represents the output of the decision tree used in the machine learning classification task, as shown in Figure 5 it is easily interpretable for individuals with experience in the field. However, employing more complex algorithms, such as neural networks, introduces complexity in both data and model explainability (Ali et al., 2023).

In Figures 4 and 5, it becomes evident that there is a necessity for incorporating an external Explainable Artificial Intelligence (XAI) model to comprehensively understand Machine Learning (ML) and Artificial Intelligence (AI) models.

5.3 XAI Algorithm Development

The XAI development stage is crucial for the XAI model functionality. During this stage, I primarily utilized the LIME package. LIME is a model-agnostic technique, applicable to any machine learning model. It aims to comprehend a model by perturbing input data samples and analysing how predictions change (Clement et al., 2023), which is essential for understanding how each instance is classified.

This is pivotal as it reveals the factors and the weight of the classification influencing a classification which can provide a more in-depth understanding of why a prediction was made or which variables influenced it without modifying the core machine learning algorithm.

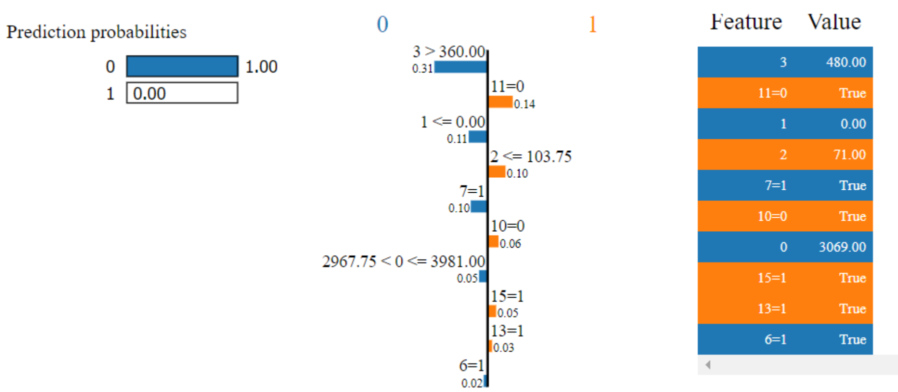


Figure 6 XAI LIME Visualisation Output

Shown in this diagram is an output of one local instance where an individual was classified as loan declined, as shown above, the LIME output indicates the feature that was used the value it contains and how it factors into the decision of whether an individual is accepted a loan.

However, the current presentation is not user-friendly and imposes a high cognitive load for comprehension. Consequently, the LIME feature-based scores have been extracted and organized into a tabular format, which will be utilized in our radar chart design.

5.4 Visualisation Design

To develop the visualization of the radar chart, the feature-based scores from the LIME output need to be transformed into a structured format containing the feature name and its corresponding weight for the classification output.

The initial approach involved using Python programming to export the visualization, employing matplotlib. However, as depicted in Figure 7, the resulting visualization was not suitable for integration into the empirical survey due to cluttered annotations and scores.

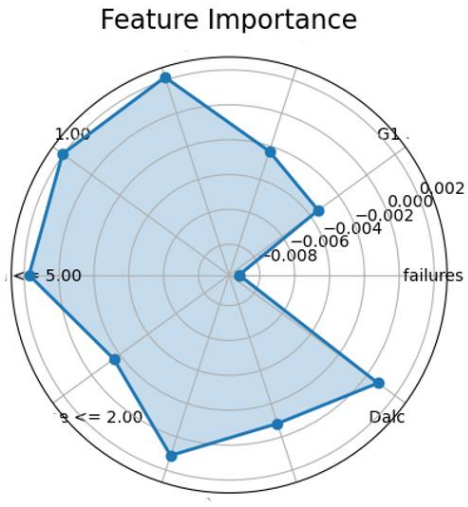


Figure 7 Radar Chart Output Matplotlib

Based on Figure 7, our original interpretation shows that traditional radar charts were visualized with a lot of interpretations or variables to consider. As a result, this may cause a high extrinsic cognitive load for the end user, given the numerous variables and data points to consider. Xiaolin et al. (2023) have suggested that increased extraneous cognitive load negatively impacts individuals' understanding and learning experience.

Furthermore, according to Seide, Jensen, and Kieser (2021), in their design of the radar chart for visualizing simulation and estimation results in network meta-analysis, they have also created a tailor-made radar chart. This demonstrates that traditional radar charts can be customized and adapted to specific end users.

Therefore, to tailor the visualization for a low cognitive load design, an external platform with greater customizability was used. In this case, the data from the XAI feature-based scores was exported to Excel for manual customization to present the radar chart more appropriately and test the customisations and features that could be adjusted.

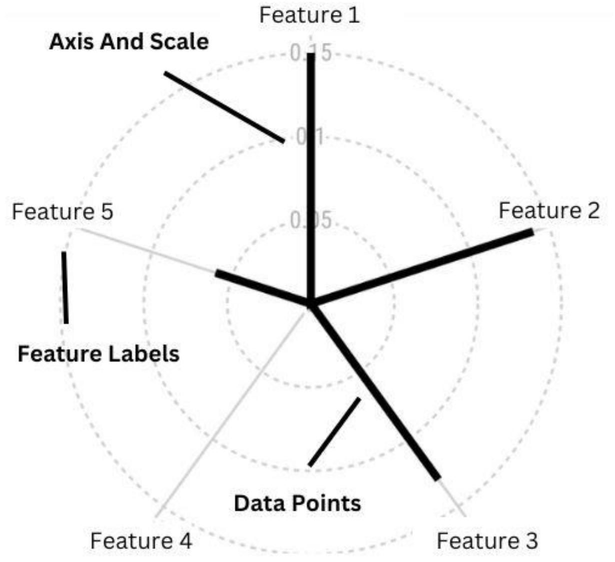


Figure 8 Customised Radar Chart Design

Figure 8, is a modified radar chart developed in Excel that has its visualization stripped down to its fundamental components.

We hypothesise that this visualisation can offer several advantages for visualising the feature attributes compared to the standard radar chart developed previously and compared to the LIME output. One of the main reasons for this is that this visualisation focuses on representing only the feature and its data with equidistant spokes or axes, each representing a different variable, with the data points plotted along these axes.

Each axis on the chart can represent a different aspect of the model's behaviour and features it utilizes to determine its classification, offering a holistic view that goes beyond simple frequency weight metrics or traditional visualizations.

Moreover, radar charts can be particularly effective in highlighting trade-offs between different factors. For instance, in our loan approval dataset, one axis might represent Income feature weight, while another represents employment status. A radar chart can vividly illustrate the compromise between these two factors, helping stakeholders understand the trade-off or the key factors that affect their loan approval.

Overall, we decided to choose this design to be studied within our empirical study, as we believe the resulting shape formed of this radar chart can provide a quick and intuitive overview of the relationships and patterns influencing the machine classification result. As a result, this design may help end-users grasp the multifaceted nature of the model's decision-making process in an optimum cognitive load.

5.5 Survey Design

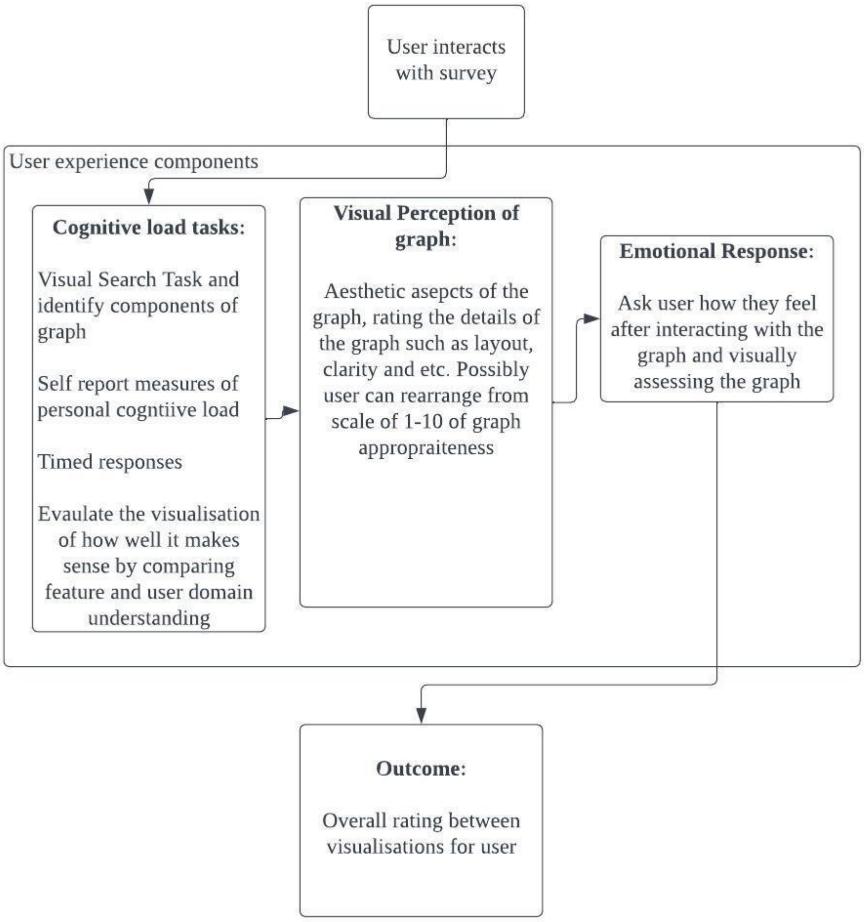


Figure 9 Flow Chart Design of Empirical Survey

Figure 9 illustrates the components and primary tasks that will be implemented to assess the cognitive load of the explanations launched for end users.

Part of the design was influenced by Huang et al.'s (2009) study on cognitive load and measuring graph visualizations. In their research, they asserted that mental effort can be measured in three ways: subjective measure (self-report), performance measure, and psychological measure.

The chosen design in this context is aligned with the subjective measure category because the other two methods require external output, making them impractical within the project's scope.

As a result, the survey developed for this honours project exclusively focuses on the subjective measure of survey participants. However, due to the subjective nature of the survey, there could be potential bias involved, but launching the survey can help as it serves as a foundational tool for understanding end users' perceptions of current explanations involved within XAI and the effectiveness of the radar chart developed.

As we could not focus on the depth of the survey due to the lack of resources available, we opted to conduct a sparse empirical survey to gain a holistic overview of the cognitive load induced on end users from the chosen explanations.

As a result, there were three main surveys launched simultaneously for this honours project, this was to measure the cognitive load of the radar chart developed, traditional LIME explanation and annotated explanation.

The hypothesis for this empirical survey is that radar charts and annotated explanations can provide a better understanding of the XAI output with optimum cognitive load. Furthermore, annotated explanation was conducted in this survey as it was used as a benchmark to measure language model explanations.

To gauge the understanding of the end user, the survey utilised the Likert methodology benchmarking. According to popular cognitive load papers such as Huang et al, (2009) in their research, they have utilised Likert scales to measure subjective cognitive load to measure graph visualisations.

As a result, using this methodology was deemed acceptable where users can rate on a scale of 1 – 5 on their interpretation of the design, where 1 is the lowest score and 10 is the highest score, in this case, the cognitive load the user interprets.

In terms of the goal of the survey, its main purpose was to understand three things from the end user, this was mainly the following:

- How does the Radar chart compare in terms of cognitive load with the XAI visualisation and annotation?
- Whether or not the Radar chart would be a better alternative to visualise feature-based data for non-technical end users.
- Whether annotation generated by a machine learning model would be an appropriate addition to help with interpretability.?

To appropriately answer these questions, the survey utilised a questionnaire methodology similar to Ouwehand et al, (2021) paper where they utilised a self-measure for end users for them to rate their own interoperability and cognitive load using a subjective rating scale.

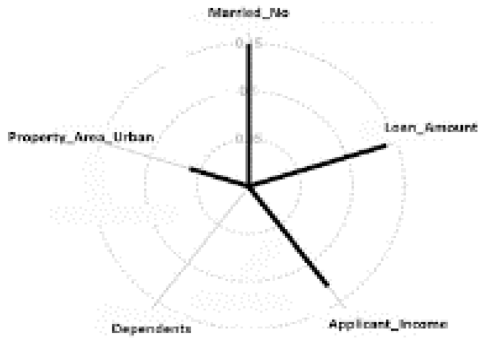
Having said that the three surveys launched had questions asking the user to conduct a visual search task for example "What is the value of the feature weight of income loan" which was then followed by "How easy was it for you to find this value?" and other feeling, emotional response and informativeness questions to gauge the end user's cognitive load.

The final survey was primarily launched online using Google survey forms which respondents were found through university classmates and online social media platforms. Furthermore, online paid platforms such as Survey Monkey were also utilised using their free trial plan in an attempt to gain more respondents to the survey.

In this section you will be answering some questions about the radar chart design please follow carefully.

Radar chart will visualize the features in a circular fashion, each line will represents the feature weight, the higher the weight the more important and vice versa.

This Diagram Below Represents A Radar Chart That classified An Individual Of Having Their Loan Approved



What Was the Lowest Used Feature? *

- ◯ Married_No
- ◯ Loan_Amount
- ◯ Applicant_Income
- ◯ Dependents
- ◯ Property_Area_Urban

How Mentally Demanding Did you Find The Task? (1 Being the easiest, whereas 5 being the hardest) *

	1	2	3	4	5	
Very Easy	◯	◯	◯	◯	◯	Very Hard

Figure 10 Survey Questionnaire for Radar Chart

5.6 Discussion of The Results

Overall, in the visualization and survey development phase, it has been demonstrated that building a tailored visualization to explain XAI outputs is feasible. However, creating the visualization using Python proved to be challenging, as it did not appropriately display the features through Matplotlib.

Therefore, in the web development phase, there requires more research to appropriately connect the backend with the frontend visualization to seamlessly generate the radar chart explanation. Research into other external modules or the creation of tailored radar charts could help streamline the process of visualizing the radar graph.

In terms of the survey development, the results were collated and cleaned using Excel spreadsheets, where unfinished or spoofed responses were excluded from the results. In total, the survey had a total of 43 properly answered participants.

Each respondent was only allowed to complete one survey out of the three. This was done to prevent cannibalization of their responses, as exposure to other surveys could influence their answers, causing bias.

The survey was designed to capture responses from a diverse range of individuals, including those with limited technical knowledge in the field.

Furthermore, considering the relatively small niche of explainable artificial intelligence, it is improbable that the participants had prior experience or knowledge in working with these systems. Thus, an assumption was made that the end users had no technical knowledge of this domain.

Although the survey was simple in terms of its design and did not capture the depth that we wanted, such as conducting more in-depth problem-solving tasks. The result of this survey answered three pivotal questions regarding the research, these were:

1. The perceived cognitive load when working with radar charts for XAI-based explanation

2. The effectiveness of radar charts for enhancing understandability of XAI output

3. Comparison of cognitive load of radar chart compared to other visualisation techniques.

Survey Question	Not effective	Slightly effective	Somewhat effective	Very effective	Extremely effective
Effectiveness of finding the desired output					
Text based response	1	3	4	3	3
LIME Bar Chart	2	2	3	5	1
Radar Chart	0	3	3	5	5

Survey Question	Very Low	Low	Moderate	High	Very High
How mentally demanding did you find the task					
Text based response	0	2	2	5	5
LIME Bar Chart	0	0	2	4	7
Radar Chart	2	3	1	6	4
What was the frustration or annoyance level while completing the survey					
Text based response	1	4	2	3	4
LIME Bar Chart	0	1	0	6	6
Radar Chart	2	3	4	4	3
How appropriate you found the visual (text/graph) for the task.					
Text based response	1	3	0	4	6
LIME Bar Chart	4	2	3	2	2
Radar Chart	2	3	2	6	3
Perception of the visual appearance in terms of clarity, attractiveness, and overall appeal					
Text based response	2	8	3	1	0
LIME Bar Chart	3	7	1	2	0
Radar Chart	1	3	5	5	2

Table 5 Result of Empirical Survey

Based on the findings presented in Table 5, the radar chart was perceived as less mentally demanding than the LIME output bar chart, though more demanding than a text-based response. In the radar chart survey, approximately 31% of participants rated it with a very low to moderate cognitive load, while only 15% of respondents in the LIME chart survey considered it to have a moderate cognitive load. Moreover, participants expressed dissatisfaction with the LIME bar chart, with 92% rating its annoyance level as high to very high. In contrast, only 43% of radar chart respondents found the annoyance level to be high to very high.

In terms of usability, respondents overwhelmingly found the radar chart effective and visually appealing. Sixty-three per cent rated it as effective to very effective, surpassing the LIME bar chart (46%) and text-based explanations (43%). Regarding visual appeal, 75% of radar chart respondents rated it as moderate to very high.

Overall, the radar chart outperformed the LIME bar chart in user perception, according to Huang et al. (2009) one of the main reasons for this could be due to the familiarity that the radar chart offers and its ease of design as it can play a crucial role in interpretability and cognitive load.

As a result, due to the overall success of the radar chart and annotation explanations, the simple empirical survey launched has solidified the original hypothesis of this study. Thus, in the web development stage, a more improved radar chart design that leverages the radar design principle combined with a language model to generate explanations will be implemented to demonstrate its real-world impact.

6. Web Application Design

In terms of the web application, its main purpose is to demonstrate how the radar chart and language model explanation can be implemented to describe an Explainable Artificial Intelligence (XAI) output for non-technical end users in a real-world application.

In the context of this honours project, a mock client named “XAI Bank” will be used as an example of a real-world banking scenario. The main problem in this context is to design a prototype XAI platform that can be used by banking customers to identify the reasons for the decline of their loans.

To design the web application, both the backend and the frontend need to be separately designed for it to function appropriately.

To further explain how this design will work, figure 11 illustrates the XAI bank website prototype, as shown there are two main pages involved in this application, the homepage followed by the XAI application page.

The XAI application page is where the end user enters their values; by doing so, these values are pushed into the back end of the application, triggering the decision tree to determine classification, which then uses the LIME XAI model to generate feature-based scores, which are used to generate the radar chart graph, and the language model will then provide an explanation.

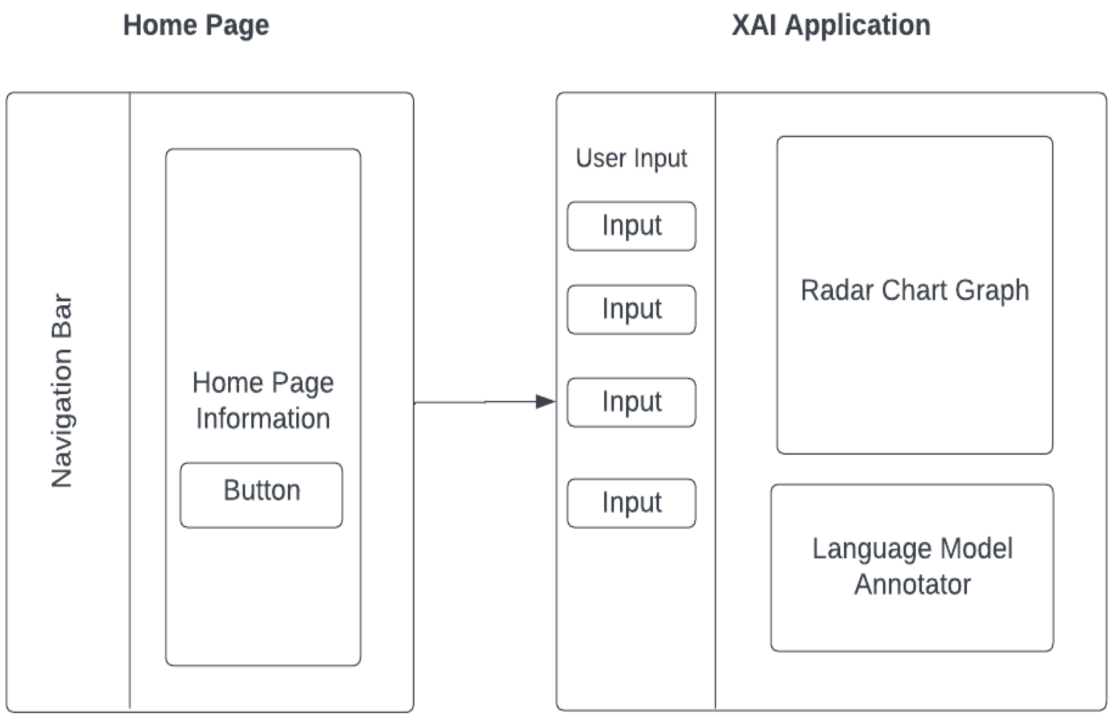


Figure 11 XAI Bank Website Prototype

To successfully develop this application, it is essential to systematically break down its key components and outline the technology required for each.

In table 6 it shows the main components that will be used for the application development.

Component	Technology Utilised	Explanation
Web Application Technology	Streamlit	Streamlit is easily deployed as a web application for machine learning projects
User Input	Streamlit Sidebar slider Streamlit Sidebar Selectbox	The slider and select box were used as interactive input for end users to provide their data to be pushed into the XAI platform.
Back End Machine Learning and XAI Algorithm	SkLearn LIME	SkLearn and LIME package was utilised for the machine learning algorithm development with extra feature engineering to extract the data
Radar Chart	Matplotlib Numpy	To output a radar chart for the web application matplotlib combined with numpy was used to make a tailored made visualisation.
Language Model Explainer	LLMchain GPT4ALL	GPT4ALL provides a pre-trained language model combining this with LLMchain which provides functionality around the language model and helps create annotation and explanation for the output of the XAI program.

Table 6 Tech Stack Table of Web Application

6.1 Machine Learning Algorithm and XAI Development

For the development of the machine learning algorithm and Explainable AI (XAI), the same methodology was utilised in the survey and visualization development phases. However, in this case, the variables for predicting the final output of the decision tree and the XAI LIME output were consolidated into a function. This was done to ensure flexibility for usage in other parts of the web app.

Additionally, a modification was made to the output of the XAI LIME features. Since LIME only outputs feature-based scores depending on the classification, they can be confusing for end users without an understanding of this domain. For instance, if a loan was accepted due to a high income, the chart might show income as a “high score” within the radar chart. Conversely, if a loan was declined due to an individual having a "Low Income," the low income would receive a “high feature” score. This inconsistency could lead to confusion for the end user.

To address this issue, the XAI bank application design primarily focuses on applicant approval. As a result, an additional feature engineering step was implemented, where I inverted the values of the features if the loan was declined so the visualisation is congruent for both loan approval and declining. This methodology aligns with the gestalt design principle (Kobourov, Mchedlidze, & Vonessen, 2015), ensuring congruity in design choices.

6.2 Web Application Radar Chart Design

The radar chart was dynamically generated using Matplotlib and Streamlit libraries and is built into a function called "plot_feature_importance." This function takes in the feature values of the eXplainable Artificial Intelligence (XAI) output in a structured format, pairing feature names with their respective scores.

The radar chart design, follows the same principle as the original empirical survey chart, however with added improvements where the features are sorted based on their importance scores, and each feature is represented by a bar in the radar chart to help with congruity.

Furthermore, the colour of each bar reflects the magnitude of the feature's importance through the use of a colour map where orange indicates positive feature importance, while purple signifies low or negative feature importance that negatively impacts an individual's success in their application.

As the bars within the radar chart were colour-coded based on the weight of the feature-based score, the statistical data points representing the feature score were improved to improve visual comprehension and reduce cognitive load.

The resulting chart not only highlights the significance of each feature but also provides a user-friendly interface for interpreting the XAI output within our web application.

Figure 12 Web APP Radar Chart Design

6.3 Language Model Design

As shown in the empirical survey, annotations can be helpful to provide necessary information for the end user to fully understand the visualisation and the understanding of each feature and it affects the overall application.

There were two main annotations generated for the web application, the first being a binary annotation. This was developed using a simple if else statement. Where if an individual has their application approved it will display:

'Your application has been APPROVED, below the orange-coloured bars have helped you receive this loan.'

Else it will output a statement:

'We apologize, your application has been DECLINED, you might want to look at the features highlighted in dark purple.'

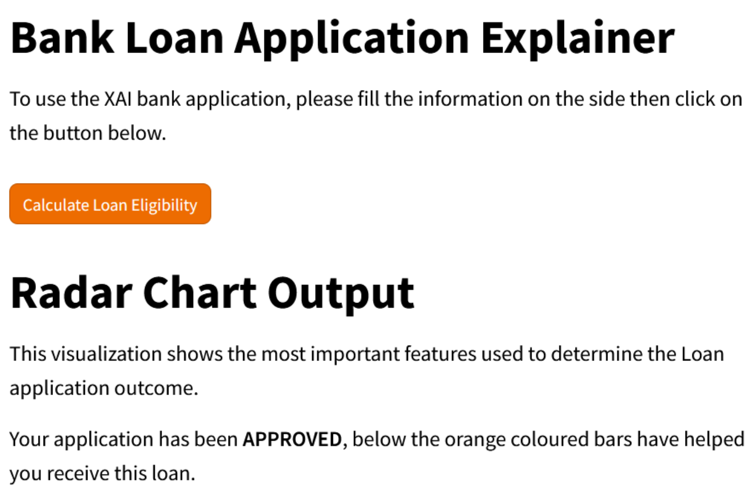


Figure 13 Web App Output of Approved Application

Given the multiple input values and combinations fed into the machine learning model, generating binary classification outputs lacks the depth needed to provide a holistic overview of their application. As a result, a language model was incorporated as an additional tool. In this project, the Orca Mini language model from GPT4ALL was integrated into the web app through LLMChain. Orca mini was chosen due to its low storage space requirements compared to other large language models and speed.

For the language model to function effectively, it requires to be run locally on my computer and have input data to be provided in its prompt. In this specific case, the feature-based score and feature name were included in the prompt and also the context of the web application to provide a simple prompt template for the LLMchain to work with, in doing so this allowed the model to inquire about the reasons behind the customer receiving a particular outcome, the features utilized, and their impact.

The loan was approved based on the features used in your application. These features included your marital status, income level, amount of the loan requested, gender, education level, self-employment status, and the location of the property where the loan is being applied for. Your application had certain criteria that were evaluated to determine if you met the requirements for a loan. The orange colored bars in your application helped to identify these criteria and ultimately led to your approval for the loan. As an individual seeking a loan, it's important to understand the features used in your application and how they impact your eligibility for a loan. This can help you make informed decisions about whether or not to apply for a loan and what terms and conditions to negotiate with lenders.

Figure 14 Language Model Output Explanation Using LLMChain

As shown in Figure 14, it produced a simple explanation of the radar chart in Figure 12 and the features involved in the application decision process. Looking at the sentence structure and explanation in Figure 14, we can see that the Orca Mini can provide end users with a simple explanation.

6.4 Final Design of Web Application

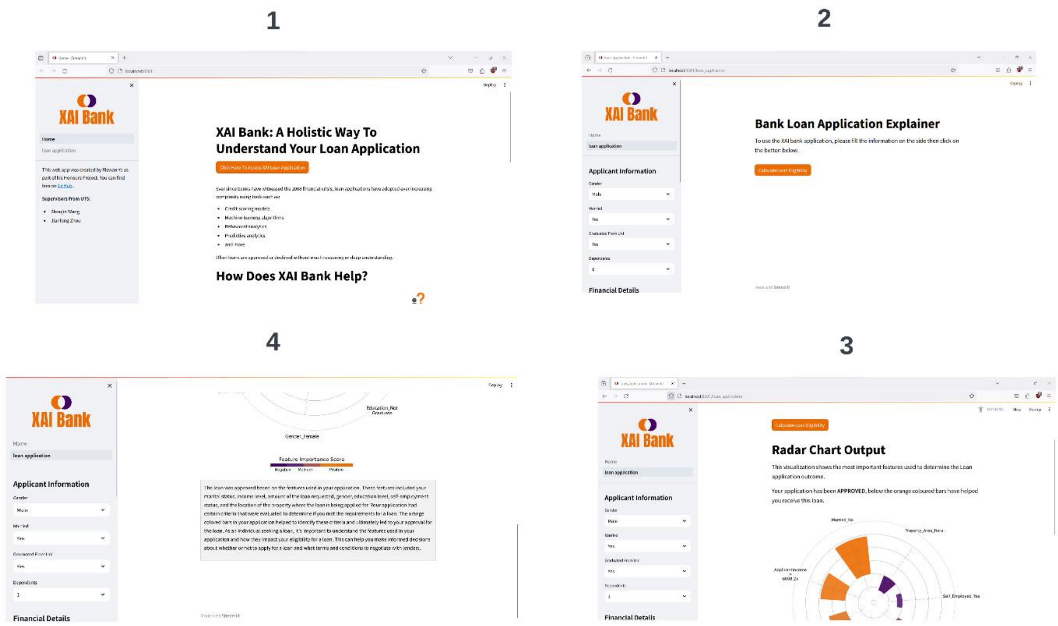


Figure 15 Overview of Final Web Application

In figure 15 it visualises the final design of the XAI bank web application. For the final design, customisations such as favicon, colour scheme, logo, loading bar, text formatting and static text explanations were added to the web application to enhance its user experience.

Looking at the figure demonstrates the step-by-step process in which the user interacts with the web app, in step 1 the user will click to enter the application, in step 2 they will input the values in the sidebar to push into the backend model and then click on run which will generate the radar chart firstly and afterwards the language model explanation, this process can be repeated if the user decides to change their inputs.

7. Discussion And Future Improvements

Based on the empirical survey and the prototype web application, we have found that radar charts combined with language model explanations can be helpful for non-technical end users to understand the decision-making processes of AI/ML programs.

The radar chart design was chosen for two specific reasons in mind, namely it is commonly used across industries to visualise feature-based data types and it is highly customisable with plenty of upgradeability which can be beneficial for model agnostic applications that are commonly used across XAI developments.

In combining cognitive load research with the gestalt principle of congruity and continuity, the final radar chart design as depicted in Figure 12 improves upon the traditional radar chart in a way that it organises features based on its weight, the area of the radar chart is removed and transferred into bars and utilises colour schemes to highlight the intensity of the values, which can provide an optimum cognitive load for end users to grasp the outcome of their application.

Furthermore, as the radar chart was limited in its capacity to fully provide context of the outcome, using a language model it was able to help supplement the shortcomings of radar charts.

However, one of the major shortcomings with radar charts is that they cannot effectively visualise outcomes that have many features involved, as it will lead to a cluttered visualisation, in the real world this can be troublesome as most deep learning models utilise many features to determine a classification.

Furthermore, in terms of the language model deployed for the prototype, it can suffer from issues such as hallucinations or generalisations as a result a more improved language model is required to provide a proper explanation.

Trust is important in the banking industry and cognitive load is a correlation to user satisfaction. Although radar chart visualisations are limited, we believe radar charts can help with non-technical end users, based on Shaw, Lee-Partridge, and Ang's (2003) research on end-user dissatisfaction, they have demonstrated that expectation of technology can impact the end users' experience and their interaction with support staff. Thus, having said that we believe a radar chart combined with a language model can help with transparency and improve cognitive load which as a result can be beneficial for industries such as the banking industry to help with customer interactions and improving the acceptance or implementation of AI/ML within society.

8. Summary & Conclusion

This study sheds light on an often-overlooked aspect of Explainable Artificial Intelligence (XAI) which is the visual analytics component for the output of these programs. The research emphasizes the importance of considering end users without a technical background in XAI. This study has revealed that radar charts emerged as the most promising and beneficial explanation for end users compared to annotation and traditional XAI outputs.

By using multidisciplinary research in combining cognitive load research with visual analytics and XAI, it has demonstrated that individuals lacking technical expertise may struggle to comprehend AI/ML outputs without easily understandable explanations. This deficiency can significantly impact user satisfaction and perception of the AI/ML model. Thus, making end users hesitant to utilise AI/ML machines.

Moreover, the study reveals that employing language models to generate information can enhance end users' interpretability of AI/ML models. However, due to limitations in research, the development of the backend machine learning algorithm and language model had sacrificed in accuracy as a result it may not represent the full extent of the program.

In conclusion, the report highlights the multifaceted nature of interpretability in AI/ML outputs. By incorporating cognitive load research with visual analytics, it has been demonstrated that XAI outputs can be further improved in transparency for the end users. Furthermore, the implementation of a web application that generates an explanation of an XAI algorithm through the use of a tailor-made radar chart and language model, demonstrates the feasibility of implementing an XAI algorithm in the real-world environment.

9. References

Alicioglu, G., & Sun, B. (2021). A survey of visual analytics for Explainable Artificial Intelligence methods. Computers & Graphics. https://doi.org/10.1016/j.cag.2021.09.002

Abeynayake, H. I. M. M., Goonetilleke, R. S., Wijeweera, A., & Reischl, U. (2023). Efficacy of information extraction from bar, line, circular, bubble and radar graphs. *Applied Ergonomics*, *109*, 103996.

Alaa, M. (2021). Artificial intelligence: Explainability, ethical issues and bias. Annals of Robotics and Automation, 034–037. https://doi.org/10.17352/ara.000011

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., ... & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion, 99, 101805.

Arya, V., Bellamy, R. K. E., Chen, P.-Y., Dhurandhar, A., Hind, M., Hoffman, S. C., Houde, S., Liao, Q. V., Luss, R., Mojsilović, A., Mourad, S., Pedemonte, P., Raghavendra, R., Richards, J., Sattigeri, P., Shanmugam, K., Singh, M., Varshney, K. R., Wei, D., & Zhang, Y. (2022). AI Explainability 360: Impact and Design. Proceedings of the AAAI Conference on Artificial Intelligence, 36(11), 12651–12657. https://doi.org/10.1609/aaai.v36i11.21540

Caron, M. S. (2019). The transformative effect of AI on the banking industry. *Banking & Finance Law Review*, *34*(2), 169-214.

Clement, T., Kemmerzell, N., Abdelaal, M., & Amberg, M. (2023). XAIR: A Systematic Metareview of Explainable AI (XAI) Aligned to the Software Development Process. Machine Learning and Knowledge Extraction, 5(1), 78–108. https://doi.org/10.3390/make5010006

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. Instructional Science, 38(2), 105–134. https://doi.org/10.1007/s11251-009-9110-0

Huang, W., Eades, P., & Hong, S.-H. (2009). Measuring Effectiveness of Graph Visualizations: A Cognitive Load Perspective. Information Visualization, 8(3), 139–152. https://doi.org/10.1057/ivs.2009.10

Kobourov, S. G., Mchedlidze, T., & Vonessen, L. (2015). Gestalt principles in graph drawing. In Graph Drawing and Network Visualization: 23rd International Symposium, GD 2015, Los Angeles, CA, USA, September 24-26, 2015, Revised Selected Papers 23 (pp. 558-560). Springer International Publishing.

Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., & Baum, K. (2021). What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. Artificial Intelligence, 296, 103473. https://doi.org/10.1016/j.artint.2021.103473

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy, 23(1), 18. https://doi.org/10.3390/e23010018

Morales-Silva, D., McPherson, C. S., Pineda-Villavicencio, G., & Atchison, R. (2020). Using radar plots for performance benchmarking at patient and hospital levels using an Australian orthopaedics dataset. Health Informatics Journal, 26(3), 2119–2137. https://doi.org/10.1177/1460458219895190

Ouwehand, K., Kroef, A. van der, Wong, J., & Paas, F. (2021). Measuring Cognitive Load: Are There More Valid Alternatives to Likert Rating Scales? Frontiers in Education, 6. https://doi.org/10.3389/feduc.2021.702616

Saeed, W., & Omlin, C. (2023). Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. Knowledge-Based Systems, 263, 110273. https://doi.org/10.1016/j.knosys.2023.110273

Seide, S. E., Jensen, K., & Kieser, M. (2021). Utilizing radar graphs in the visualization of simulation and estimation results in network meta-analysis. *Research Synthesis Methods*, *12*(1), 96-105.

Shankar, R., & Duraisamy, S. (2018). Visual Analytics Need, Process, Scope, Tools and Techniques and Challenges. International Journal of Computer Sciences and Engineering, 6(10), 153–158. https://doi.org/10.26438/ijcse/v6i10.153158

Shao Xiaolin, Suwarsi Suwarsi, Pornman Inta, Asmarani Lathu Fajarina, Muflih Muflih, Listyana Natalia Retnaningsih, & Santi Damayanti. (2023). The Relationship Between Learning Concentration and Understanding Level Through the Online Learning Process. Jurnal Keperawatan Respati Yogyakarta, 10(2), 89–89. https://doi.org/10.35842/jkry.v10i2.741

Shaw, N., Lee-Partridge, J. E., & Ang, J. S. (2003). Understanding the hidden dissatisfaction of users toward end-user computing. *Journal of Organizational and End User Computing (JOEUC)*, *15*(2), 1-22.

Skulmowski, A., & Xu, K. M. (2021). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational psychology review*, 1-26.

Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. Information Fusion, 76, 89–106. https://doi.org/10.1016/j.inffus.2021.05.009

Zhang, Q., Hall, M., Johansen, M., Galetic, V., Grange, J., Quintana-Amate, S., Nottle, A., Jones, D. M., & Morgan, P. L. (2022). Towards an Integrated Evaluation Framework for XAI: An Experimental Study. Procedia Computer Science, 207, 3884–3893. https://doi.org/10.1016/j.procs.2022.09.450