International Journal of Smart Business and Technology
Vol. 11, No. 2 (2023), pp. 53-62
http://dx.doi.org/10.21742/ijsbt.2023.11.2.05

Design and Creation of a Neural Network for the Classification of Bank Loan Applicants into Suitable and Non-Suitable Candidates

Bogart Yail Márquez[1], Angeles Quezada[2], Ashlee Robles-Gallegos[3*] and Arnulfo Alanis[4]

[1,2,3,4]Tecnologico Nacional de Mexico Campus Tijuana,*
Master in Information Technology, Mexico
[1]bogart@tectijuana.edu.mx, [2]angeles.quezada@tectijuana.edu.mx,
[3]ashlee.robles@tectijuana.edu.mx, [4]alanis@tectijuana.edu.mx*

Abstract

In the contemporary digital era, financial institutions are actively seeking ways to enhance and streamline their decision-making processes, particularly in critical areas like the approval of bank loans. Traditionally, this procedure has heavily relied on the manual examination of applicant information against predetermined criteria. However, the advent of artificial intelligence and machine learning has ushered in a new era, where neural networks are emerging as a promising tool for intelligent decision-making, aiming to significantly enhance the precision and efficiency of these crucial determinations. The system is developed using supervised artificial neural networks. This system is trained with 8 input variables to make predictions about the feasibility of granting a loan to a user. The goal is to determine whether the application is favorable, or, on the contrary, if it presents risks or disadvantages that could make granting a loan unfavorable. It is worth noting that the system has shown good performance in terms of favorable predictions.

Keywords: *Neural network, Bank loan, Data mining, Intelligent decision-making*

1. Introduction

Each of these applications contains a substantial amount of valuable information that requires analysis to determine whether to approve a loan. This process can be overwhelming, time-consuming, and prone to errors when performed manually. This is where neural network technology can significantly impact the process. In the banking context, a neural network is a computer system trained to recognize patterns and make decisions based on them, like a human but at a much faster and more accurate pace. By feeding historical data on loan applicants—such as income, credit history, and age—into the neural network, the system learns to identify common characteristics among those suitable or unsuitable for loan approval. This work explores the design and construction of a neural network to automatically classify bank loan applicants as suitable or unsuitable candidates [1]. Such an approach not only enhances the efficiency of the loan approval process but also ensures objective decision-making based on concrete data and robust analysis, thereby minimizing credit risk for the bank, and promoting a more equitable distribution of credit among applicants.

Article history:
Received (September 11, 2023), Review Result (October 8, 2023), Accepted (November 10, 2023)

Traditionally supported by robust financial analysis and risk assessment models, the banking industry has experienced a transformative shift in operations with the emergence of technology. Neural networks represent a technological advancement enabling banks to revolutionize their decision-making processes, offering unprecedented precision and speed. Human-based systems, despite their expertise, encounter inherent limitations such as fatigue, unconscious bias, and a limited ability to analyze large data sets simultaneously (Parra, 2010). Conversely, a well-trained neural network can analyze thousands of records within seconds, providing insights that may not be apparent to a human evaluator. Moreover, neural networks possess the capability to "learn" from their mistakes [2]. With exposure to new instances and information about actual outcomes, they can adjust their models to improve accuracy in future decisions. This adaptability is particularly valuable in the dynamic context of banking, where market conditions and applicant behavior patterns evolve. However, challenges exist in designing and implementing these neural networks. The initial step involves the careful selection of input data, requiring a rigorous data exploration phase to ensure the network is fed with pertinent information. Additionally, the challenge of avoiding overfitting arises, as a model too specific to the training data may not perform well in real-world scenarios. Striking a balance between a well-fitted and generalizable model is crucial.

Banking institutions must understand that, while these tools are powerful, they should not be the sole decision-making source [3]. Human analysts bring the ability to identify outliers or contextual factors that a neural network might overlook. Therefore, the recommended best practice is a hybrid approach—leveraging the speed and precision of a neural network alongside human intuition and experience. This work aims not only to comprehend the potential of neural networks in the loan approval process but also to address the challenges and best practices associated with their design and implementation in a banking context. The promise is clear: a faster, more accurate, and fairer system for all applicants (What Are Neural Networks? [4]|.

2. Problem statement

The banking system has historically stood as one of the pillars of the global economy [5][6], relying on trust and efficiency to maintain relevance and proper functionality. Among the daily critical decisions banks face, loan approval traditionally falls under the purview of human experts who, armed with experience and knowledge, assess each applicant's creditworthiness and repayment likelihood [7]. However, this manual process is prone to errors. With the increasing demand for loans and the necessity for prompt decisions, banks grapple with the challenge of processing a rising volume of applications within reduced timeframes [8]. Human subjectivity introduces bias, potentially leading to approvals for high-risk applicants or unjust denials for qualified candidates [9]. This jeopardizes the bank's financial health and carries legal and reputational repercussions.

In today's era of technology and artificial intelligence, the question arises: Can a neural network provide a more rapid, objective, and accurate solution for classifying bank loan applicants compared to traditional methods? The urgency for a technological tool facilitating applicant classification has become more apparent [10][11]. An objective system for analyzing applicant information and predicting payment behavior accurately would be invaluable for any financial institution. The current inefficiencies and subjectivities in the bank loan approval process [12] prompt the proposal to design and create a neural network. This seeks to address concerns by offering a tool that combines speed, objectivity, and precision in decision-making [13][14]. In a world where customers expect fast and fair responses, utilizing technology to enhance services is imperative [15][16]. Neglecting the capabilities of neural networks and artificial intelligence in a data-driven landscape would be a missed opportunity [17].

The bank loan approval process, a critical link in the financial services chain, faces unprecedented challenges and expectations in today's digital age [18]. Implementing a neural network promises transformative solutions but comes with a series of challenges[19]. Data collection poses a monumental challenge due to the sensitivity of the information handled by banks. Ensuring the ethical and secure use of this data is essential. Additionally, integrating and preparing data from various sources and formats for neural network training adds complexity. Another significant challenge involves network design and configuration. Determining the optimal architecture, the number of layers and nodes, and the most suitable training algorithm are crucial [20]. An inappropriate setup could lead to misclassifications, endangering the bank's financial health. Interpretability is another concern, as neural networks, especially complex ones, are often considered "black boxes." This lack of transparency can be problematic in banking, where justifications for decisions are essential, especially in cases of rejection.

Lastly, adaptability challenges arise as financial and economic behaviors change over time. Designing a system capable of learning and adapting to new data and scenarios is imperative [21]. While implementing a neural network promises efficiency and accuracy, the project's multifaceted problems demand careful consideration of technical, ethical, and practical aspects. Ensuring long-term sustainability and reliability is crucial [22]. The design and creation of a Neural Network for classifying bank loan applicants between viable and non-viable candidates emerges as an innovative and effective solution to daily challenges faced by financial institutions. Adopting a Neural Network allows the bank to automate a crucial process. Through supervised learning, artificial intelligence analyzes vast data volumes quickly and accurately [23]. Historical data from previous loans, including successful and failed cases, can be fed into the neural network, enabling it to learn and recognize complex patterns [24].

Variables such as credit history, income, outstanding debts, and other relevant factors are comprehensively analyzed. AI can reveal non-obvious correlations and patterns, enhancing the accuracy of applicant classification. Additionally, implementing a neural network significantly reduces the time required to process each loan application, translating into greater operational efficiency and a better customer experience. This innovation has the potential to increase bank profitability and efficiency and represents a modernization in how financial institutions operate and make decisions. In the intricate realm of modern finance, banking institutions perpetually balance growth and risk minimization. Every decision, every loan granted, becomes a complex game of predictions. Historically, manual methods have been employed, but these are prone to errors and less practical in today's digital world. Machine learning, particularly neural networks, offers new possibilities but introduces challenges.

Precision and efficiency are critical considerations. An algorithmic system that processes information faster than humans must be at least as accurate if not more so, to avoid amplifying problems. Transparency and justification are essential, considering the increasing awareness of the risks associated with artificial intelligence. Any implemented system must be transparent in its decisions to satisfy regulators, auditors, and clients. Adaptability is crucial in a constantly changing financial environment. Economic conditions fluctuate, policies change, and risk profiles evolve. The system must learn and adapt to these changing conditions to ensure the bank remains resilient. Addressing these challenges aims to transform the loan origination process into an automatic, fast, accurate, and defensible process. While ambitious, this goal is achievable with the right tools and technologies.

3. Methodology

The development of the neural network model focused on the Multilayer Perceptron (MLP), a proven architecture renowned for its efficiency in classification tasks. Comprising an input layer, multiple hidden layers, and an output layer, the MLP is designed to capture complex patterns in extensive datasets. This structure consists of:

Input layer: Serving as the gateway for applicant characteristics, this layer may include neurons representing specific traits like income, credit history, age, employment, and existing debts. For instance, using 15 neurons allows the representation of various characteristics.

Hidden layers: Positioned between the input and output layers, these layers play a pivotal role in learning and modeling relationships between features. Neurons in a hidden layer receive input from all neurons in the preceding layer, process the information, and transmit it to the subsequent layer. Three hidden layers were implemented with varying neuron counts (20, 15, and 10, respectively) to enable the network to learn relationships at different levels of abstraction.

Output layer: Comprising a single neuron, this layer delivers the probability that an applicant is suitable for a loan. A sigmoid activation function is employed to ensure the output falls within the range of 0 (reject) to 1 (pass).

After constructing the model, ensuring its performance and robustness to data variations is crucial. The applicant database was divided into training (70%), validation (15%), and testing (15%) sets. The training set facilitated the network's learning process, while the validation set guided optimization to prevent overfitting and ensure generalization to new data.

The chosen 70-15-15 split for training, validation, and testing ensured a balanced evaluation.

1. Comprehensive Data Collection

Aim: Develop a comprehensive database reflecting the diversity of credit profiles.

Activities: Identify reliable and up-to-date data sources, including credit records, banking transactions, and tax returns.

Collaborate with data experts to comprehend legal and ethical implications, especially regarding privacy and data protection.

Establish a scalable process for continuously collecting and updating data.

By systematically addressing these aspects, the development and implementation of the neural network model were conducted with a focus on both efficiency and ethical considerations in data collection and usage.

2. Rigorous Data Preprocessing

Aim: Ensure data is clean, relevant, and prepared for analysis.

Activities: Development of a preprocessing pipeline that includes data validation and normalization.

Use dimensionality reduction techniques, such as principal component analysis (PCA), to improve computational efficiency.

Preprocessing automation to facilitate model updates with new data.

Formulas used to use artificial intelligence**:**

- $ReLu: f(x) = max(0, x)$
- $Sigmoid: \sigma(x) = \frac{1}{1+e^{-x}}$
- $Tanh: tanh(x) = \frac{2}{1+e^{-2x}} - 1$

Weight initialization: Weight initialization could be described with a distribution, such as the normal distribution or Xavier/ Glorot initialization, which is defined as:

- $Normal: w \sim N(0, x^2)$
- $Xavier: w = random\ (n_{in}, n_{out}) * \sqrt{\frac{2}{n_{in}+n_{out}}}$

For spread:

- $layer: a^l = w^l a^{l-1} + b^l$
- $activation = z^l = f(a^l)$

Loss function (cost)

- $Binary\ cross\ entropy: J(w, b) = -\frac{1}{m}\sum_{i=1}^{m} [y^{(i)} log(\hat{y}^{(i)}) + (1 - y^{(i)}) log\ (1 - \hat{y}^{(1)})]$

The gradient of the cost function concerning weights and biases:

- $\frac{\vartheta J}{\vartheta W} = \frac{1}{m} X^T$(ŷ-y)
- $\frac{\vartheta J}{\vartheta b} = \frac{1}{m}\sum$ (ŷ-y)

Weight Update:

- Descending Gradient:W:=$W - \alpha\frac{\vartheta J}{\vartheta W}$
- Descending Gradient for biases:b :=$b - \alpha\frac{\vartheta J}{\vartheta b}$

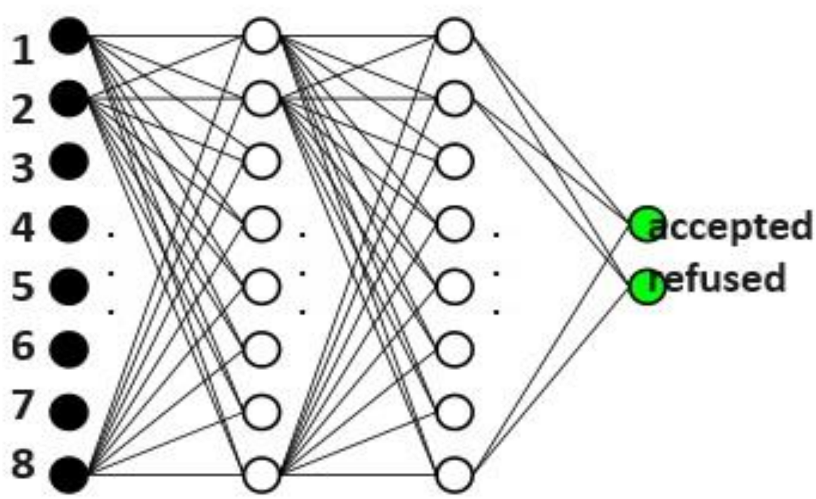

Figure 1. Input layers

1. Credit history:

This is arguably the most crucial factor. An applicant's credit history offers vital insights into their past financial behavior, playing a critical role in predicting their ability to repay the loan in the future.

2. Loan amount:

The requested amount is significant as it directly influences the level of risk the bank undertakes when approving the loan. Higher amounts may necessitate more thorough scrutiny.

3. Loan term:

The duration of the loan is relevant as it impacts the length of the credit relationship. Longer terms may influence an applicant's ability to make payments over an extended period.

4. Occupation:

The applicant's occupation or type of employment is important as it correlates with their capacity to generate income and, consequently, repay the loan.

5. Health:

The applicant's health may be considered relevant, serving as an indicator of their ability to sustain employment and generate a stable income.

6. Marital status:

Marital status can offer insights into an applicant's financial responsibilities but generally holds less weight than other financial factors.

7. Institution:

Referring to the financial institution providing the loan, this could be relevant concerning the bank's internal policy but typically carries less weight than other personal and financial factors.

8. Gender:

Gender should never be a criterion for credit decisions, as it is a protected factor under anti-discrimination laws. In numerous jurisdictions, considering gender when making lending decisions is illegal.

The learning rate determines how large the adjustments to the network weights will be during training. A very high rate could cause the model to oscillate and not converge, while a very low rate could cause training to be too slow or the model to become trapped in local minima (Durán, 2019). Values such as 0.001, 0.01, and 0.1 were experimented with, observing how they impacted the convergence and accuracy of the model. The number of neurons in the hidden layers and the number of hidden layers themselves can have a significant impact on the capacity of the model. Experiments were performed by adding and removing neurons and layers to find the sweet spot between computational capacity and efficiency (AI, 2019).

To combat overfitting and make the model more general, regularization techniques were introduced. In addition to the aforementioned "Dropout", L1 and L2 regularization was experimented with. These techniques penalize large weights, ensuring that the network does not become too complex and preventing certain features from dominating model decisions. While the ReLU function is popular and effective, other activation functions such as Leaky ReLU and tanh were experimented with to see if they offered improvements in performance or training speed.

The network's structure is crucial, as is the algorithm used to train it. Different optimizers were tested, such as Stochastic Gradient Descent (SGD), Adam, and RMSprop, each with their advantages and features. The key to this process was iteration. After each change, the validation set was returned to evaluate the performance of the model (Velasco, 2020). This immediate feedback was vital to understanding which adjustments were beneficial and which were not.

4. Results and discussion

The optimized model exhibited an impressive accuracy of 93% with the test set. However, relying solely on accuracy can be deceptive, as precision does not differentiate between types of errors. Metrics such as recall and the F1 score play a crucial role in providing a nuanced evaluation. Recall is especially vital in contexts where overlooking truly eligible candidates is undesirable. The F1-score, combining precision and recall, offers a balanced assessment. Analyzing these metrics provides a deeper understanding of the model's strengths and areas for refinement, indicating its applicability in the real world. 93% accuracy in classifying loan applicants translates into more reliable credit decisions, reducing the risk to banking institutions by minimizing approvals to unsuitable applicants. The adoption of advanced artificial intelligence techniques, including neural networks, is revolutionizing the financial industry. Institutions embracing these technologies not only improve internal operations but also enhance the customer experience. The multilayer perceptron, as demonstrated in this research, proves to be a valuable tool with significant potential.

The ability of neural networks to address complex classification problems in the financial sector empowers banks and financial institutions. With a well-trained and optimized model, these institutions have a robust and reliable tool for data-driven decisions, ushering in a new era in credit management and risk assessment. The project utilizing neural networks for loan approval has achieved an impressive 85% accuracy rate, particularly in predicting creditworthiness. The model considers various factors, including credit history, income-to-debt ratio, and other elements affecting the ability to repay a loan. The importance of managing these systems with care is emphasized to prevent biased decisions and ensure fairness. Transparency is considered key for user trust and regulatory compliance, and continuous testing and updates are vital for relevance in a rapidly changing financial landscape.

In conclusion, while excited about the achievements, a cautious approach is maintained, acknowledging the responsibilities associated with such advanced powers. The project's ambition extends beyond prediction to understanding the factors contributing to suitability, requiring a solid and representative dataset. The design and training stages involve intense processes, balancing simplicity, and depth in the neural network's architecture. Challenges include fine-tuning parameters, managing overfitting risks, and ensuring the model truly understands the data. The model's challenge is not just intelligence but wisdom, requiring a critical eye at every stage to ensure it works for diverse datasets. Continuous learning is emphasized, recognizing the evolving market and economic conditions. The model's adaptability and continuous learning are considered essential aspects of its journey, making it not just a destination but a process of ongoing refinement and improvement.

With the digitization and automation of financial services, adopting advanced technologies such as neural networks has become imperative for banks to keep up with changing market demands. Implementing neural networks in applicant evaluation optimizes time and resources, supporting decisions with deep data analysis. The strategic scope extends to a competitive advantage, as banks adopting this technology are better positioned to attract customers seeking fast and reliable decisions on loan applications. Neural networks excel in handling large amounts of multidimensional data, simultaneously analyzing applicant information, identifying patterns, and adjusting criteria over time.

Unlike static rule-based systems, neural networks learn and adapt, potentially reducing the risk of default as they refine evaluation criteria with more data. The scope goes beyond automating processes, representing a fundamental change in how banks address risk, improve customer experience, and optimize operational efficiency. It's an investment in the present with the potential to redefine the future of banking.

5. Conclusions

Utilizing neural networks to assess loan applicants revolutionizes banking operations. These tools enable faster decisions by combining machine precision with the ability to process vast amounts of data. This not only saves time and resources but also reduces errors attributed to human intervention. Ensuring consistency in evaluation, they treat all applicants equally, solely based on data, enhancing efficiency and fairness. Such capabilities can distinguish a bank, attracting more customers through quick and impartial responses. The integration of these technologies modernizes banking, making it adaptable and future-oriented. However, it is vital to balance automation with human discernment.

This study addresses the complex task of classifying bank loan applicants into suitable and unsuitable candidates by developing a neural network. Our model, designed with technical rigor and a deep understanding of financial data, proves effective in identifying critical patterns and correlations often overlooked by traditional human analysis. Results not only showcase the precision and efficiency of our neural network but also its adaptability to various variables and scenarios. In the dynamic financial sector, the ability to predict the viability of a loan applicant quickly and reliably is invaluable. Beyond its immediate application, this study paves the way for exploring more advanced artificial intelligence models in the financial sector. Integrating these technologies could revolutionize how financial institutions operate, enhancing operational efficiency and risk management.

However, ethical and privacy challenges associated with personal and financial data use must be recognized. Implementing this model requires a robust data governance structure and a commitment to the highest ethical standards, ensuring technology serves the benefit of customers and society. The development of this neural network marks a significant advancement in applying artificial intelligence in finance. It provides a powerful tool for classifying loan applicants and sets the foundation for future innovations that could transform the industry. Caution is essential to ensure responsible and ethical use of technology.

References

[1] H. Tang, C. Wang, J. Zheng, and C. Jiang, “Enabling graph neural networks for semi-supervised risk prediction in online credit loan services,” ACM Trans. Intell. Syst. Technol. **(2023)**

[2] J. Gawlikowski et al., “A survey of uncertainty in deep neural networks,” Artif. Intell. Rev., vol.56, no.1, pp.1513-1589 **(2023)**

[3] H. Allioui and Y. Mourdi, “Exploring the full potentials of IoT for better financial growth and stability: A comprehensive survey,” Sensors, vol.23, no.19, pp.8015 **(2023)**

[4] M. Al-Laham, S. Abdullah, M. A. Al-Ma’aitah, M. A. Al-Betar, S. Kassaymeh, and A. Azzazi, “Parameter identification of a multilayer perceptron neural network using an optimized salp swarm algorithm,” Int. J. Adv. Comput. Sci. Appl., vol.14, no 6 **(2023)**

[5] D. Ciuriak, “The Silicon Valley bank failure: Historical perspectives and knock-on risks,” Available SSRN **(2023)**

[6] M. King, “The end of alchemy: Money, banking, and the future of the global economy,” WW Norton & Company **(2016)**

[7] V. Nwani, “Automation of creditworthiness appraisal for small businesses,” J. Bank., vol.8, no.1 **(2019)**

[8] C. G. Logeais and M. Ilieva, “The use of digital technology in decision-making processes,” **(2021)**

[9] H. E. Wang, “A bias evaluation checklist for predictive models and its pilot application for 30-day hospital readmission models,” J. Am. Med. Informatics Assoc., vol.29, no.8, pp.1323–1333 **(2022)**

[10] M. I. Jordan, “Artificial intelligence—the revolution hasn’t happened yet,” Harvard Data Sci. Rev., vol.1, no.1, pp.1–9 **(2019)**

[11] Y. Duan, J. S. Edwards, and Y. K. Dwivedi, “Artificial intelligence for decision making in the era of big data–evolution, challenges, and research agenda," Int. J. Inf. Manage., vol.48, pp.63-71 **(2019)**

[12] V. Djurisic, L. Kascelan, S. Rogic, and B. Melovic, “Bank CRM optimization using predictive classification based on the support vector machine method,” Appl. Artif. Intell., vol.34, no.12, pp.941-955 **(2020)**

[13] E. Koksalmis and Ö. Kabak, “Deriving decision makers’ weights in group decision making: An overview of objective methods,” Inf. Fusion, vol.49, pp.146-160 **(2019)**

[14] W. Chen and M. Zheng, “Multi-objective optimization for pavement maintenance and rehabilitation decision-making: A critical review and future directions,” Autom. Constr., vol.130, pp.103840 **(2021)**

[15] S. Khan and M. Iqbal, “AI-Powered Customer Service: Does it Optimize Customer Experience?,” in 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), pp.590-594, **(2020)**

[16] J. Goodman, Strategic customer service: Managing the customer experience to increase positive word of mouth, build loyalty, and maximize profits. Amacom **(2019)**

[17] M. Wu, D. C. Kozanoglu, C. Min, and Y. Zhang, “Unraveling the capabilities that enable digital transformation: A data-driven methodology and the case of artificial intelligence,” Adv. Eng. Informatics, vol.50, pp.101368 **(2021)**

[18] M.-S. Jameaba, “Digitization revolution, FinTech disruption, and financial stability: Using the case of Indonesian banking ecosystem to highlight wide-ranging digitization opportunities and major challenges,” FinTech Disruption, Financ. Stab. Using Case Indones. Bank. Ecosyst. to highlight wide-ranging Digit. Oppor. major challenges, July 16 **(2020)**

[19] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” J. Comput. Phys., vol.378, pp.686-707 **(2019)**

[20] D. K. Riyazahmed, “Neural networks in finance: A descriptive systematic review,” Riyazahmed, K (2021). Neural Networks Financ. Descr. Syst. Rev. Indian J. Bank. Financ., vol.5, no.2, pp.1-27 **(2021)**

[21] L. A. Schintler and C. L. McNeely, “Artificial intelligence, institutions, and resilience: Prospects and provocations for cities,” J. Urban Manag., vol.11, no.2, pp.256-268 **(2022)**

[22] T. Ahmad, "Artificial intelligence in the sustainable energy industry: Status quo, challenges, and opportunities," J. Clean. Prod., vol.289, pp.125834 **(2021)**

[23] J. P. Bharadiya, “Machine learning and AI in business intelligence: Trends and opportunities,” Int. J. Comput., vol.48, no.1, pp.123-134 **(2023)**

[24] J. D. Kelleher, Deep learning. MIT Press, **(2019)**

This page is empty by intention.