Anomaly Detection on Dynamic Data Streams using Continual Learning

Submitted in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

Submitted by

Pratham Gupta (IIT2020026)

Shashwat Kumar Gautam (IIT2020030)

Jay Suthar (IIT2020087)

With the supervision of

Dr. Manish Kumar

(Associate Professor)

Department of Information Technology

Indian Institute of Information Technology - Allahabad

Prayagraj-211015

July 2024

CANDIDATE DECLARATION

We **(Pratham Gupta, Shashwat Gautam, Jay Suthar)**, Roll No.**(IIT2020026, IIT2020030, IIT2020087)** certify that this B.tech project report work entitled **Anomaly Detection on Dynamic Data Streams using Continual Learning** is submitted by us in partial fulfillment of the requirement of the degree of B.tech IT in the Department of Information Technology, Indian Institute of Information Technology, Allahabad.

We understand that plagiarism includes,

1. Reproducing someone else's work (fully or partially) or ideas and claiming it as one's own.
2. Reproducing someone else's work (verbatim copying or paraphrasing) without crediting.
3. Committing literary theft (copying some unique literary construct).

We have given due credit to authors/sources through proper citation for all words, ideas, diagrams, graphics, computer programs, experiments, results, websites, that are not our original contribution, We have used quotation marks to identify verbatim sentences and given credit to original authors/sources.

We affirm that no portion of our work is plagiarized. In the event of a complaint of plagiarism, we shall be fully responsible. We understand that our supervisor may not be in a position to verify that this work is not plagiarized.

Pratham Gupta (IIT2020026)
Shashwat Kumar Gautam (IIT2020030)
Jay Suthar (IIT2020087)
Department of Information Technology
Indian Institute of Information Technology - Allahabad

CERTIFICATE FROM SUPERVISOR

It is certified that the work contained in the B.tech project report titled **Anomaly Detection on Dynamic Data Streams using Continual Learning** by **Pratham Gupta, Shashwat K. Gautam, and Jay Suthar** has been carried out under my supervision and that this work will not be submitted elsewhere for a degree.

Signature of the Supervisor: ———————————-

Name of the Supervisor: **Dr. Manish Kumar**

Designation: **Associate Professor**

Department: **Dept. of Information Technology**

Indian Institute of Information Technology - Allahabad

CERTIFICATE OF APPROVAL

The foregoing thesis is hereby approved as a creditable study in Information Technology and its allied areas. It is carried out and presented in a satisfactory manner to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but the thesis only for the purpose for which it is submitted.

Committee Members for Evaluation of the Thesis for Final Examination:

.....

.....

.....

.....

ACKNOWLEDGEMENTS

First of all, we would like to thank our institute **Indian Institute of Information Technology, Allahabad**. This institute gave a unique platform to learn new techniques and enhance our technical skills. We are highly grateful to the honorable Director, **IIIT Allahabad, Prof. Mukul Sharad Sutaone** for his helping attitude and encouragement to excel in studies.

We extend our thanks to our brilliant thesis supervisor, **Dr. Manish Kumar, Associate Professor, IIIT Allahabad**, for his excellent ideas and insights. This report would not have been possible without his many important contributions, and we are heartily grateful for his open and honest feedback. Having a good thesis supervisor is key when studying for a higher degree, and we could simply not have wished for a better guide and work partner than **Dr. Manish Kumar**.

We would like to express our deepest appreciation towards Mr.Manish Kumar Maurya, Research Scholar IIIT Allahabad. We would also like to thank our family members for their valuable moral support in our life.

Pratham Gupta, Shashwat K. Gautam, Jay Suthar

IIT2020026, IIT2020030, IIT2020087

B.Tech (IT)

Abstract

Anomaly detection in dynamic data streams poses a significant challenge for conventional machine learning techniques due to evolving data distributions and concept drift. To address this challenge, this research project focuses on enhancing anomaly detection methods by leveraging continual learning approaches. Continual learning enables models to adapt to changing data patterns while retaining previous knowledge, thus mitigating false positives and improving accuracy. The project aims to develop models capable of accommodating new information over time while maintaining their ability to detect anomalies accurately. By integrating real-time adaptation mechanisms, the models can continuously update their understanding of normal and abnormal behavior, ensuring robust performance in dynamic environments. The primary objective is to contribute to the development of more effective and adaptive anomaly detection systems for real-world finance data streams. Through continual learning techniques, the research seeks to empower anomaly detection models to evolve alongside dynamic data streams, addressing challenges such as stationary training setups and catastrophic forgetting. The proposed methodology emphasizes improving accuracy and reducing false positive rates, critical factors in anomaly detection systems' reliability. By evaluating and refining these techniques through rigorous experimentation, this research aims to establish a foundation for practical deployment in financial scenarios, offering enhanced performance and reliability in detecting anomalies amidst evolving data distributions and concept drift.

Contents

Abstract **v**

Contents **vi**

List of Figures **viii**

List of Tables **ix**

Abbreviations **x**

1 Introduction **1**
1.1 Continual Learning 4
1.1.1 Continual Learning over Traditional ML 5

2 Literature survey **7**

3 Problem definition and Objectives **17**

4 Methodology **18**
4.1 Continual Anomaly Detection Framework 18
4.1.1 Autoencoder Networks (AENs) for Anomaly Detection Setup . 18
4.1.2 Working of AENs with CL 22
4.1.3 Continual Learning Setup 24
4.2 Datasets 28
4.3 Comprehensive Examination of Data 29
4.3.1 Identification of Anomalous Entries 30
4.4 Preprocessing 31
4.5 Experimental Setup 35
4.6 Performance Metrics 37
4.6.1 Precision 37
4.6.2 Recall 38
4.6.3 F1 Score 38

5 Results and Discussion **39**

6 Conclusion and Future Scope **43**

Bibliography **45**

List of Figures

4.1 AENs Architecture [1] 19

5.1 Average Reconstruction Loss For Each Experience in DA and DB . . 40
5.2 Average Reconstruction Loss For Each Experience in DC 40

List of Tables

2.1 Literature Survey Part-1. 9
2.2 Literature Survey Part-2 10
2.3 Literature Survey Part-3. 11
2.4 Literature Survey Part-4. 12
2.5 Literature Survey Part-5. 13
2.6 Literature Survey Part-6. 14
2.7 Literature Survey Part-7. 15
2.8 Literature Survey Part-8. 16

5.1 Avg. L_{Rec} For Each Experience 39
5.2 Evaluation Metrics for the Injected Anomalies 41

Abbreviations

CL	Continual Learning
ML	Machine Learning
DL	Deep Learning
AENs	Autoencoder Networks
GEM	Gradient Episodic Memory
SI	Synaptic Intelligence
MAS	Memory Aware Synapses
L_{Rec}	Reconstruction Error

Chapter 1

Introduction

Since the early 20th century, the fields of Data Science, AI, ML, and DL have received significant attention across various industries. Recently, with the emergence of technologies like IoT and advancements in decision-making systems, Artificial Intelligence has become a fundamental aspect of this century. The rapid expansion of computing power and database capacities has driven forward technologies. However, within the abundance of normal data entries, anomalies or outliers, that significantly deviate from the norm, present a formidable challenge. Anomaly detection is a vital task in data mining with extensive applications [2]. Although various definitions exist for anomalies, they all converge on the idea of data points that diverge from typical patterns. This thesis investigates the significance of anomaly detection and explores its diverse definitions, shedding light on its critical role across various domains and majorly focus on the anomaly detection related to financial auditing data.

The detection of fraudulent activities within financial accounting data relies on identifying unusual patterns present in journal entries. Individuals aiming to benefit personally through the misuse of organizational resources often deviate from typical posting behaviors, resulting in abnormal attribute values in a small portion of journal entries. Auditors utilize various Computer Assisted Audit Techniques (CAATs)

to uncover these anomalies, including rule-based analyses and statistical methods. Rule-based analyses may involve examining uncommon posting times or frequent changes in vendor bank accounts, while statistical methods like Benford's Law or extreme value analysis aid in detecting irregularities. Incorporating deep-learning into modern audit procedures poses certain challenges, particularly in the realm of anomaly detection within financial accounting data. Firstly, deep-learning models encounter the hurdle [3] of stationary training environments, where they are trained on a fixed and limited dataset of journal entries, potentially missing out on pertinent shifts in distributions. Secondly, these models are prone to Catastrophic Forgetting, where the integration of new data can disrupt previously acquired knowledge. This is especially concerning in continuous auditing scenarios where models are exposed to a steady influx of data. These difficulties highlight the intricacies involved in effectively identifying anomalies in financial accounting data using deep-learning methodologies.

To tackle the issue of erroneous outcomes in real-world audits, CL has emerged as a excellent strategy. CL, or lifelong learning, entails updating deep-learning models incrementally as new data emerges. This proves crucial in unsupervised anomaly detection within extensive accounting datasets, where changes from previous financial periods can affect audit conclusions. False positive alerts happen when the system mistakenly flags genuine activities as fraudulent, resulting in resource wastage. Conversely, false negative decisions occur when the system fails to identify actual instances of fraud, posing considerable risks. Drawing inspiration from CL's effectiveness in diverse fields, researchers are exploring its adaptation in financial audits. Their approach involves ongoing learning from real-world journal entries, showcasing enhanced performance compared to traditional re-training methods.

Within the realm of financial accounting, an anomaly signifies a departure or irregularity from anticipated patterns or standards within financial statements. These anomalies may take various shapes, including unexpected fluctuations in revenue,

unexplained shifts in expenses, inconsistencies in balance sheet figures, or peculiar trends in financial ratios. In financial accounting, anomalies can serve as indicators of errors in transaction recording or reporting, potential manipulation or fraudulent activities, alterations in business operations or market dynamics, or underlying organizational issues. Detecting anomalies holds paramount importance for upholding the accuracy and dependability of financial information and for facilitating informed decision-making processes. Financial analysts and auditors employ a range of methodologies, such as ratio analysis, trend assessment, and statistical techniques, to pinpoint anomalies in financial data. Furthermore, technological advancements, including data analytics and artificial intelligence, have empowered the development of more sophisticated anomaly detection approaches.

Anomaly detection is aimed at pinpointing patterns that diverge from the anticipated behavior within a dataset [2]. It entails the recognition of data points, occurrences, or observations that notably deviate from the norm or expected behavior.

An array of methods exists for anomaly detection, encompassing statistical approaches, machine learning algorithms, and pattern recognition techniques. These methods typically entail comparing observed data points to historical data or predefined models [4] to ascertain whether they represent normal behavior or anomalies.

Anomaly detection in financial accounting data entails uncovering irregularities, discrepancies, or uncommon patterns within financial statements and transactions. Such anomalies could signify errors, fraudulent activities, or other matters necessitating further investigation.

Various techniques and approaches are utilized to detect anomalies in financial accounting data, including:

- Ratio Analysis: Examining financial ratios over time or against industry standards to detect unusual fluctuations.

- Trend Analysis: Assessing trends and patterns in financial data to identify deviations from expected behavior.
- Statistical Methods: Utilizing statistical tests and models to detect outliers or anomalies in the data distribution.
- Machine Learning: Employing machine learning algorithms to automatically recognize patterns and anomalies in large volumes of financial data.

By implementing robust anomaly detection mechanisms, organizations can enhance their financial transparency, mitigate risks, and ensure compliance with regulatory requirements, ultimately fostering trust and confidence among stakeholders.

1.1 Continual Learning

It represents a machine learning paradigm focused on the continuous enhancement of models over time without forgetting previously acquired knowledge. Unlike traditional [5] approaches that train models on fixed datasets, continual learning systems adapt and refine their performance with each new data input. This paradigm proves particularly beneficial in environments where data distributions evolve or new information is regularly introduced. Instead of retraining from scratch, continual learning involves updating models incrementally, and incorporating fresh data while retaining past insights. Its key attributes encompass incremental updates, ensuring models adjust to changing data dynamics; knowledge retention, preventing the loss of previously learned information; adaptability to [6] evolving environments or tasks; and resource efficiency, as incremental updates often demand fewer computational resources than full retraining. Continual learning finds applications across fields such as NLP, computer vision, robotics, and anomaly detection, enabling models to sustain relevance and efficacy over time amidst shifting conditions.

1.1.1 Continual Learning over Traditional ML

Continual learning offers distinct advantages over traditional machine learning (ML) approaches, particularly in anomaly detection, by addressing two significant challenges: catastrophic forgetting and concept drift.

- Catastrophic Forgetting: Traditional machine learning models often replace previously acquired knowledge with new data during training, resulting in what is known as catastrophic forgetting. This is problematic in anomaly detection, where historical anomalies hold importance for understanding evolving patterns of fraud or errors[7]. Continual learning tackles catastrophic forgetting by allowing the model to update incrementally while retaining knowledge from previous sessions [8]. This ensures a comprehensive understanding of both historical and new anomalies, reducing the risk of catastrophic forgetting.
- Concept Drift: Concept drift arises when there's a shift in the fundamental data distribution over time, diminishing a model's efficacy in identifying anomalies in fresh data. In anomaly detection, concept drift can result from changes in business operations, market dynamics, or fraud tactics. Continual learning addresses concept drift [9] by continuously adapting the model to evolving data patterns. Instead of assuming a static distribution, continual learning models dynamically update parameters to capture changes, maintaining high detection accuracy even with concept drift.

By overcoming catastrophic forgetting and concept drift, continual learning brings several benefits to anomaly detection [10]:

- Robustness to Changes: Continual learning models are more robust to changes in data distribution, making them effective in dynamic environments.
- Long-Term Performance: These models maintain high performance over time by adapting to evolving data patterns and retaining knowledge from past anomalies.

- Improved Accuracy: Continual learning models capture subtle changes in anomalous behavior, leading to improved detection accuracy compared to traditional ML approaches.

CL offers a more adaptive and robust framework for anomaly detection, ensuring effective detection in evolving data distributions while mitigating the risks of catastrophic forgetting and concept drift.

Adapting to anomalies in dynamic data streams challenges conventional machine learning techniques. This project enhances anomaly detection in financial data by leveraging continual learning to address evolving distributions and concept drift. By enabling models to retain previous knowledge while accommodating new information, we aim to mitigate false positives and improve accuracy. Real-time adaptation mechanisms will further enhance understanding of normal and abnormal behavior, contributing to more effective anomaly detection systems for real-world finance applications.

The subsequent sections are organized as follows: Initially, we offer an overview of related work. Subsequently, we introduce the proposed continual learning framework for unsupervised anomaly detection in financial data streams. Following this, we detail the experimental setup and outline results. The work culminates in a discussion and summary, covering key insights and potential future directions.

Chapter 2

Literature survey

A thorough examination of recent research reveals a diverse landscape in anomaly detection and continual learning, particularly in the context of evolving data environments such as IoT, multivariate time series, and dynamic data streams.

Several studies address the challenge of concept drift adaptation, proposing adaptive approaches for anomaly detection in evolving data environments like IoT systems [2] [5]. These methods aim to dynamically adjust to changing data patterns, enhancing detection effectiveness. Similarly, [6] introduces DiEvD, a methodology for detecting disruptive events from dynamic data streams, with a specific focus on Twitter data. While innovative, its specificity to Twitter may limit its generalizability.

Deep learning techniques are prevalent in anomaly detection [3], particularly in multivariate time series data. These studies review various deep learning methods, architectures, and applications. However, they acknowledge limitations in empirical validation across domains and scalability issues for large datasets [5].

Continual learning is another recurring theme, with several papers surveying methods and applications [7] [8]. They cover techniques such as fine-tuning, transfer

learning [9], and feature extraction with pre-trained models. Additionally, approaches to mitigate catastrophic forgetting, a common challenge in continual learning, are discussed[10].

Autoencoders [11] play a significant role in anomaly detection, with studies exploring their architectures and applications [12]. While these papers provide a comprehensive overview, they may lack depth in certain niche applications or recent advancements[13].

Financial fraud detection is a specific domain of interest, with [14] reviewing various anomaly detection techniques and recent advances. Despite advancements, scalability and real-time detection capabilities remain challenges[15] [16]. While the literature review highlights various approaches to anomaly detection and continual learning methodologies, there is a notable absence of studies specifically focusing on applying continual learning techniques to dynamic financial data streams.

To address this gap, our research aims to develop and evaluate novel continual learning algorithms tailored to the unique characteristics of financial data streams. By leveraging continual learning, the approach seeks to adaptively update anomaly detection models to evolving market conditions, thereby enhancing detection accuracy and robustness in real-time financial environments. Furthermore, we aim to investigate the scalability and efficiency of continual learning algorithms in handling large-scale financial datasets, addressing a key challenge identified in the literature. Through empirical validation and comparative analysis, our work contributes to advancing anomaly detection capabilities in dynamic financial data streams, complementing the existing literature on continual learning and anomaly detection methodologies.

Table 2.1: Literature Survey Part-1.

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[2] Lijuan Xu, Xiao Ding, Haipeng Peng, Dawei Zhao, Xin Li	ADTCD: An adaptive anomaly detection method for IoT, tackling concept drift by dynamically adjusting to evolving data patterns for enhanced detection.	ADTCD proposes adaptive anomaly detection for IoT, autonomously monitoring data streams, detecting concept drift, and updating models to enhance robustness.	INSECTS-Abr INSECTS-Inc INSECTS-IncGrd INSECTS-IncRec SWAT WADI BATADAL	Limitations Focus on homogeneous industrial control scenario limits generalizability. Limited abnormal data coverage hampers anomaly detection effectiveness.	Explore self-supervised learning to tackle anomaly label scarcity. Extend ADTCD beyond autoencoder-based models. Improve decision-making by building diverse anomaly detectors.
[3] Gen Li, Jason J Jung	The paper explores deep learning techniques for detecting anomalies in multivariate time series, addressing applications and research challenges in this field.	The research explores deep learning in multivariate time series anomaly detection, assessing models, real-world use, limitations, and future prospects.	Various signal dataset, Yahoo dataset , Fall detection dataset	The paper lacks cross-domain empirical validation. Deep learning anomaly detection faces scalability issues with large datasets, while interpretability challenges impede anomaly understanding.	Future research aims to establish benchmark datasets for evaluating deep learning in time series anomaly detection. Explainable AI models are crucial for interpretation. Domain-specific applications offer tailored solutions.

Table 2.2: Literature Survey Part-2

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[4] Durgesh Samariya, Amit Thakkar	This paper aims to comprehensively survey anomaly detection algorithms, categorizing them by domain, and discussing their strengths, weaknesses, and applications.	The paper extensively reviews anomaly detection algorithms, categorizing them into statistical, ML, DL, ensemble, and time-series-specific approaches, detailing principles, strengths, limitations, and applications.	Not provided.	The paper's limitations include potential gaps in coverage of anomaly detection methods and a lack of empirical evaluations, as it is a survey rather than an empirical study.	Enriching the paper with recent anomaly detection advancements ensures relevance. Diverse case studies offer practical insights, while comparative analyses empower decision-making for researchers.
[5] Qiuyan Xiang, Lingling Zi, Xin Cong, Yan Wang	The paper reviews concept drift adaptation in deep learning, aiming to summarize current techniques and provide insights for addressing drift.	The paper conducts a thorough literature review on methods for adapting deep learning models to concept drift, summarizing approaches, strengths, weaknesses, and applications across domains.	KDD CUP 1999, Weather, Spam, and CoverType	The paper may lack coverage of all concept drift methods, lack empirical evaluation for some methods, and could be biased, affecting overall conclusions.	Future research could validate concept drift adaptation methods through experiments, develop new techniques addressing limitations, and evaluate methods across diverse domains for robustness.

Table 2.3: Literature Survey Part-3.

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[6] Aditi Seetha, Satyendra Singh Chouhan, Emmanuel S Pilli, Vaskar Raychoudhury	DiEvD proposes a continual machine learning method for detecting disruptive events in real-time Twitter data streams, enhancing adaptability to dynamic changes.	The approach utilizes continual machine learning to detect disruptive events from dynamic datastreams, focusing on real-time adaptation to changes in Twitter data for identifying impactful events.	DET and nDET Dataset, Twitter Dataset, Event2012	The paper's reliance on Twitter data may limit generalizability to other dynamic datastreams, constraining applicability beyond social media.	Future research may expand the proposed methodology beyond Twitter to diverse data streams like news feeds, sensors, or financial markets. Continual ML techniques could be refined for better disruptive event detection, exploring new algorithms or architectures.
[7] Anton Lee, Yaqian Zhang, Heitor Murilo Gomes, Albert Bifet, Bernhard Pfahringer	The paper introduces SurpriseNet, aiming for efficient class incremental learning inspired by anomaly detection, enabling learning new classes without replaying past data, enhancing adaptability.	The approach integrates anomaly detection techniques to handle class incremental learning challenges. SurpriseNet likely uses neural networks to adapt to new classes while mitigating retraining needs and addressing concept drift.	S-DSADS S-PAMAP2 S-FMNIST S-CIFAR10 S-CIFAR100	The review assesses scalability, generalization, robustness, and evaluation comprehensiveness of concept drift adaptation methods, crucial for their applicability across diverse datasets and domains.	The paper aims to extend the method to diverse domains, enhance efficiency for large-scale data, and address concept drift in incremental learning.

Table 2.4: Literature Survey Part-4.

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[8] Da-Wei Zhou, Hai-Long Sun, Jingyi Ning, Han-Jia Ye, De-Chuan Zhan	The paper surveys CL methods using pre-trained models, facilitating adaptation to new tasks or domains without extensive retraining, summarizing approaches, challenges, and trends in this field.	The paper reviews CL techniques with pre-trained models, encompassing methodologies like fine-tuning, transfer learning, and feature extraction, analyzing their effectiveness, scalability, and applicability across domains.	CUB Inc10, IN-R Inc5, IN-A Inc20, OmniBench Inc30	Insufficient discussion on practical challenges or real-world deployment issues. Lack of comparison or evaluation metrics for the surveyed methods.	Future research could develop continual learning methods to mitigate catastrophic forgetting. Standardized benchmarks and evaluation protocols are essential for fair comparison and reproducibility across datasets.
[9] Gen Li, Jason J JungLiyuan Wang, Xingxing Zhang, Hang Su, Jun Zhu	The paper aims to offer a thorough survey of continual learning, covering theory, methodologies, and applications. It seeks to understand and address challenges, solutions, and diverse domain applications.	The paper systematically reviews literature on continual learning, analyzing theoretical frameworks, methodologies, algorithms, and applications. It categorizes approaches, identifies challenges, and discusses advancements in the field.	Not Provided	The paper may exhibit coverage bias, lacking inclusivity of all continual learning research, potentially providing a shallow overview due to the broad topic and focusing on specific subdomains.	Future research may explore meta-learning, generative modeling, and attention mechanisms in continual learning. Cross-domain applications across computer vision, NLP, and robotics are suggested, along with standardized evaluation metrics develop

TABLE 2.5: Literature Survey Part-5.

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[10] M Mundt, Y Hong, I Pliushch, V Ramesh	The paper aims to explore continual learning in deep neural networks, addressing challenges, lessons, and gaps between continual and active/open-world learning, proposing bridging approaches.	The paper comprehensively reviews continual learning in deep neural networks, analyzing methodologies, challenges, and proposed solutions. It offers insights into addressing complexities and may introduce new frameworks.	Not Provided	The paper comprehensively reviews continual learning in deep neural networks, analyzing methodologies, challenges, and proposed solutions. It offers insights into addressing complexities and may introduce new frameworks.	Empirical validation on benchmark datasets and real-world applications, exploration of novel methodologies for addressing challenges like scalability, and interdisciplinary collaboration to integrate continual learning with related fields.
[11] Pengzhi Li, Yan Pei, Jianqiang Li	This paper aims to offer a comprehensive overview of autoencoder design and its applications within the realm of deep learning.	Conducting an exhaustive review of existing literature, the paper thoroughly examines diverse architectures and utilization scenarios of autoencoders in deep learning.	Not Provided	While comprehensive, the survey may lack depth in certain niche applications or recent advancements, given the rapid evolution of deep learning research.	Future investigations could delve into the development of novel autoencoder architectures tailored to specific tasks and evaluate their efficacy in real-world applications spanning various domains.

Table 2.6: Literature Survey Part-6.

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[12] Kamal Berahmand, Fatemeh Daneshfar, Elaheh Sadat Salehi, Yuefeng Li, Yue Xu	To provide a comprehensive survey of autoencoders and their diverse applications in machine learning.	Conducted a thorough review of literature to discuss various types of autoencoders, their architectures, and applications across different domains.	Not Provided	The survey lacks detailed analysis on the performance comparison of different autoencoder architectures and may overlook recent advancements in the field.	Future research could focus on conducting empirical studies to compare the effectiveness of various autoencoder architectures on different datasets and exploring novel applications of autoencoders in emerging domains.
[13] Jiahao Yu, Xin Gao, Feng Zhai, Baofeng Li, Bing Xue, Shiyuan Fu, Lingli Chen, Zhihang Meng	This research seeks to develop a strong method for identifying anomalies in multivariate time series data. It combines adversarial training and contrastive learning techniques to achieve this goal.	The proposed method employs an adversarial contrastive autoencoder framework to establish a latent representation space that effectively separates normal and anomalous patterns. Adversarial training is utilized to boost the discriminative capabilities of the learned representations, while contrastive learning encourages compactness within classes and separability between them.	SWaT, SMD, PSM, MSL, SMAP	The method may face challenges when dealing with highly imbalanced datasets or rare anomaly occurrences. Moreover, its training process could demand considerable computational resources due to the complexity of the adversarial and contrastive objectives.	Subsequent investigations could concentrate on mitigating imbalanced dataset issues by employing techniques like data augmentation or specialized loss functions. Additionally, exploring the scalability of the method to larger datasets and real-time applications would enhance its practical applicability.

Table 2.7: Literature Survey Part-7.

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[14] Zhongju Sun, Jian Wang, Yakun Li	To propose a novel unsupervised method, RAMFAE, for detecting anomalies in visual data using autoencoder architecture.	RAMFAE employs an autoencoder framework to reconstruct normal data, leveraging reconstruction errors to identify anomalies without labeled training data.	MNIST	RAMFAE's performance might degrade with complex datasets or varying anomaly types due to its reliance on reconstruction errors, potentially leading to false positives or negatives.	Enhancing RAMFAE's robustness to diverse anomalies by integrating additional features or refining the autoencoder architecture, and exploring its applicability across different domains with larger datasets for validation.
[15] Waleed Hilal, S. Andrew Gadsden, John Yawney	This paper aims to comprehensively review anomaly detection techniques for financial fraud, highlighting recent advances and their applicability.	Through a systematic literature review, the paper consolidates various anomaly detection methods in financial fraud detection, emphasizing recent advancements and their effectiveness.	Not Provided	Despite the advancements, some techniques may lack scalability or struggle with real-time detection, and others may have limitations in detecting sophisticated fraudulent activities.	Future research could focus on enhancing the scalability and real-time capabilities of existing techniques, integrating multiple approaches for improved accuracy, and addressing emerging challenges posed by evolving fraudulent strategies.

Table 2.8: Literature Survey Part-8.

Authors (Year)	Aim	Approach	Datasets	Shortcomings	Future Work
[16] Marco Schreyer, Timur Sattarov, Damian Borth, Andreas Dengel, Bernd Reimer	Designing a technique to identify irregularities within extensive accounting datasets using deep autoencoder networks.	Utilized deep autoencoder networks to learn features from accounting data and detect anomalies by reconstructing normal data patterns.	SAP ERP	Limited exploration of interpretability of detected anomalies, potential challenges in scalability to extremely large datasets, and sensitivity to hyperparameter tuning.	Investigate methods to enhance interpretability of detected anomalies, optimize the scalability of the approach for handling extremely large datasets, and explore novel techniques for automated hyperparameter tuning.

Chapter 3

Problem definition and Objectives

Adapting to anomalies in dynamic data streams presents a significant obstacle for conventional machine learning techniques. These streams constantly evolve, resulting in shifting data distributions and concept drift. Static models struggle to keep pace with these changes, leading to elevated rates of false positives and diminished accuracy in anomaly detection. This undermines the efficacy of anomaly detection systems, highlighting the pressing need for approaches capable of handling evolving data patterns without succumbing to catastrophic forgetting.

This project aims to enhance anomaly detection methods on dynamic financial data streams using continual learning techniques. By leveraging continual learning, we aim to develop models capable of adapting to evolving data distributions and concept drift, addressing the challenges of stationary training setups and catastrophic forgetting. Our objective is to improve anomaly detection performance by enabling models to accommodate new knowledge over time while retaining previously learned information. Additionally, we will focus on real-time adaptation mechanisms to continuously update models and improve their understanding of normal and abnormal behavior. Ultimately, our goal is to contribute to the development of more effective and adaptive anomaly detection systems for real-world applications in finance and beyond, ensuring accurate detection while mitigating false positive rates.

Chapter 4

Methodology

4.1 Continual Anomaly Detection Framework

Our Continual Anomaly Detection Framework integrates autoencoders for anomaly detection, continual learning techniques, and the adaptation of AENs-based anomaly Detection within the continual learning paradigm. This approach enables robust anomaly detection while accommodating evolving data distributions.

4.1.1 Autoencoder Networks (AENs) for Anomaly Detection Setup

Autoencoders stand as a pivotal archetype within the realm of neural networks, serving as proficient learners capable of compressing and reconstructing data [11]. This group of neural networks stands out for their proficiency in unsupervised learning scenarios, where input and output coincide, and the network aims to simplify and condense the data of input into a more refined representation.

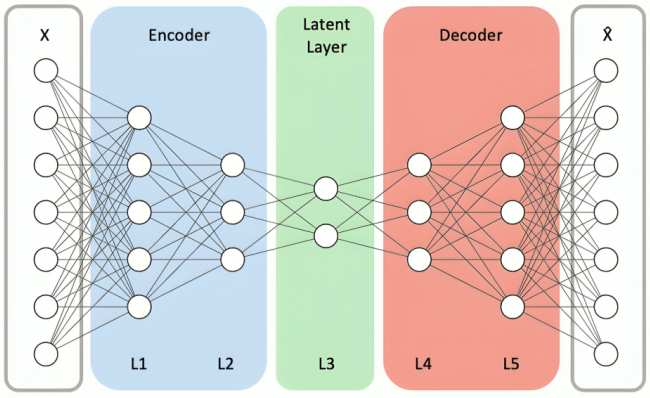

FIGURE 4.1: AENs Architecture [1]

- **Components of Autoencoder Networks (AENs):**
  - Encoder: The encoder, located at the front of the network, compresses input data into a lower-dimensional representation, commonly known as the bottleneck or latent representation[11]. Through various layers of neurons, the encoder distills the essence of the input data, abstracting away redundant or less significant information.
  - Decoder: The decoder, which complements the encoder, aims to reconstruct the initial input data from the compressed representation created in the bottleneck. Employing symmetrical layers to the encoder, the decoder strives to reverse the compression process, reinstating the essential details of the input data.
- **Working of Autoencoder Networks (AENs):** Autoencoder Networks (AENs) function by acquiring the ability to encode and then reconstruct input data, with the objective of reducing the disparity between the original input and the reconstructed output. [13]. Here's how AENs typically work:
  - **Encoder Phase**: In the encoder phase, the input, such as attributes of journal entries, undergoes a sequence of transformations within the neural network across multiple layers. Each layer applies mathematical operations to reduce the input's dimensionality. Through this process, the network abstracts the input, learning to extract and capture its significant

features and patterns [16]. These features are gradually distilled and synthesized into a compressed representation known as the latent space or latent representation.

The latent space encapsulates the essential characteristics of the input data in a more compact and expressive form. The encoder phase condenses the information from journal entries into the latent space, preserving essential attributes and patterns while minimizing redundancy and irrelevant details. This compressed representation serves as a distilled essence of the original data, facilitating efficient processing and analysis in subsequent stages of the autoencoder architecture.

Let the encoder network, be denoted by q_ψ , a function parameterized by ψ , which maps the input data x_i (e.g., a journal entry) from dimensional space $\mathbb{R}^k$ to latent space $\mathbb{R}^m$ where $k > m$. Mathematically, the encoder can be represented as $z_i = q_\psi(x_i)$ where z_i is the latent representation (or code vector) of the input x_i.

– **Decoder Phase**: During the decoder phase, the latent representation generated by the encoder undergoes a reversal of the compression process. This step involves feeding the latent representation through multiple layers in the neural network. With each layer, the dimensionality of the latent space increases gradually.

As the latent representation traverses through the decoder layers, it undergoes transformations aimed at reconstructing the original input data. These transformations work in tandem to regenerate the attributes of the journal entries, striving to replicate them as faithfully as possible.

The final output of the decoder phase is the reconstructed version of the original input data. This reconstructed data endeavors to capture the essence of the journal entries, aiming to mirror their attributes with precision [16]. The decoder phase of an autoencoder fulfills its primary

goal of compressing and accurately reconstructing the original information by reconstructing the input data. This process allows the network to efficiently learn representations of intricate data.

Let the decoder network, be denoted by p_ϕ, a function parameterized by ϕ, which maps the latent representation z_i back to the original high-dimensional space $\mathbb{R}^k$ attempting to reconstruct the data from input. Mathematically, the decoder can be represented as $\tilde{x}_i = p_\phi(z_i)$ where $\tilde{x}_i$ is the reconstruction of the input x_i. based on its latent representation z_i.

– **Training**: In the training phase, the Autoencoder Network (AEN) undergoes training using a dataset consisting of normal, non-anomalous journal entries. The main goal of training is for the model to reduce the difference between the input and output data. This minimization typically occurs through the optimization of a chosen loss function, such as mean squared error (MSE) or binary cross-entropy.

Throughout the training process, the parameters of the neural network, including weights and biases, are iteratively adjusted to reduce the reconstruction error [14]. This adjustment aims to ensure that the output produced by the autoencoder closely resembles the input data provided during training. By fine-tuning the network's parameters, the autoencoder learns to encode and decode the input data effectively, capturing the essential features and patterns present in the journal entries.

Mathematically, the Autoencoder Network (AEN) aims to reduce the difference between the original input data x_i and its reconstruction $\tilde{x}_i$ where $\tilde{x}_i = p_\phi(q_\psi(x_i))$. The objective is to find the best model parameters $\theta^* = \{\psi^* \cup \phi^*\}$ that minimise L_{Rec} across all input data points $i = 1, 2, ...N$.

$$L_{Rec} = \arg\min_{\phi,\psi} \|x_i - \tilde{x}_i\| \tag{4.1}$$

– **Anomaly Detection**: Once the AEN has been trained, it can be utilized for identifying anomalies by calculating the reconstruction error for each input data point [14]. Entries with high reconstruction errors are flagged as potential anomalies, indicating deviations from the normal patterns learned during training.

  In a real-world audit scenario, entries with high reconstruction errors may undergo further examination by auditors to determine if they represent genuine anomalies or errors.

  This paradigm offers a promising avenue for detecting anomalies within financial accounting data, leveraging the power of neural networks to discern subtle deviations from regular patterns. As we delve deeper into the realm of anomaly detection using continual learning, the foundation laid by AENs serves as a cornerstone in our endeavor to fortify the audit process against irregularities and discrepancies.

  In essence, the AEN leverages its learned understanding of normal data patterns to discern anomalies within unfamiliar data. By training exclusively on normal data and pinpointing deviations during inference, AENs serve as adept tools for detecting anomalies in financial accounting datasets.

4.1.2 Working of AENs with CL

Autoencoder Networks (AENs) within a Continual Learning (CL) framework adjust to changing data distributions and assimilate knowledge from an ongoing flow of journal entry data. In this setup, AENs operate by continually refining their representations of input data, enabling adaptation to evolving information patterns over time.

- **Incremental Learning:**
  - In a continual learning setup, the AEN model undergoes continuous training on a sequence of separate experiences or batches of data. Each experience represents a subset of journal entry data collected over a specific time period, such as a day or week.
  - As new experiences are observed, the model incrementally learns from the data without forgetting previously learned information. This ensures that the AEN remains up-to-date and adapts to changes in the financial accounting environment.
- **Adaptive Training:**
  - The AEN undergoes adaptive training on each new experience, updating its parameters (weights and biases) based on the data encountered in that experience. This allows the model to capture any shifts or trends in the data distribution over time.
  - Adaptive training helps the AEN maintain its ability to reconstruct normal, non-anomalous journal entries accurately, even as the characteristics of the data evolve.
- **Continual Improvement:**
  - With every fresh encounter, the AEN (Automated Entry Network) enhances its understanding of the inherent structures and connections within the data found in journal entries. This continual learning process enables the model to enhance its anomaly detection capabilities over time.
  - By continually refining its understanding of normal data patterns and anomalies, the AEN becomes more adept at distinguishing between regular and irregular journal entries, thereby improving the accuracy of anomaly detection.

- **Retention of Previous Knowledge:**
  - Despite learning from new experiences, the AEN retains knowledge learned from past experiences. This is achieved through techniques such as parameter regularization or rehearsal-based methods, which help mitigate the risk of catastrophic forgetting.
  - Retaining previous knowledge ensures that the AEN can effectively detect anomalies based on historical data patterns, while also adapting to new trends or anomalies as they emerge.

4.1.3 Continual Learning Setup

In the domain of CL, we embark on a journey to enhance the anomaly detection capabilities within In this scenario, a complex model labeled as f_θ, such as an autoencoder represented as $f_\theta : q_\psi \bigoplus p_\phi$. This model functions within a dynamic environment where data is continuously observed as a series of separate N experiences denoted as $\{E_i\}_{i=1}^{N}$.

In the context of auditing, each experience E_i encapsulates a specific period of organizational posting activities, denoted as D_i originating from M distinct organizational posting activities PA_j. These posting activities collectively form the foundation of each experience E_i, shaping the data landscape that the model encounters.

Upon encountering the data stream, the model f_θ is tasked with adapting and learning from each experience E_i, refining its understanding of the underlying patterns and anomalies within the financial accounting data. This adaptive learning process leads to the emergence of experience-specific models denoted as f_θ^i where θ represents the model parameters defining its architecture and behavior.

Central to the training process is the computation of the reconstruction loss(L_{Rec}) for each experience E_i. This loss metric quantifies the disparity between the original

data distribution within PA_j and the reconstructed data outputted by the model f_θ^i. Mathematically, the reconstruction loss L_{Rec}^i for the i-th experience is expressed as:

$$L_{Rec}^i = \sum_{j=1}^{M} L_{Rec}(f_\theta^i, PA_j) \tag{4.2}$$

where PA_j represents the organizational posting activities constituting the experience E_i.

Through this continual learning framework, the model f_θ iteratively refines its anomaly detection capabilities, leveraging insights gained from each new experience to adapt and improve its performance over time. As the journey unfolds, this adaptive approach to learning holds the promise of enhancing the audit process, enabling the detection of anomalies with greater accuracy and efficacy amidst the ever-evolving landscape of financial accounting data.

In each facet of this setup, the anomaly detection model embarks on a journey of incremental learning, navigating through a stream N-1 of preceding experiences denoted as $\{E_i\}_{i=1}^{N-1}$. At each juncture, the model immerses itself in the wealth of insights gleaned from past encounters, leveraging them as stepping stones for continual refinement.

Transitioning from one experience to the next, a crucial element is the initialization of model parameters θ. Here, the model's journey is guided by the torchbearer of optimal learning, as the parameters θ are primed with the wisdom distilled from the preceding experience E_{i-1}. This strategic initialization sets the stage for the model to build upon its past learnings, harnessing them as a foundation upon which to further hone its anomaly detection prowess.

With each progressive encounter, the model dynamically adapts, assimilating new insights while retaining the essence of its prior knowledge [8]. This iterative process of incremental learning imbues the model with a nuanced understanding of

the evolving data landscape, empowering it to navigate the intricacies of financial accounting data with precision and agility.

Here we will be using 3 continual learning methods [9][8]-

- **Gradient Episodic Memory (GEM)**: This method addresses catastrophic forgetting by storing important past gradients and ensuring that the current task's gradient does not conflict with them. This enables the model to maintain its knowledge from prior tasks while acquiring new ones. In Gradient Episodic Memory (GEM), the objective is to optimize the parameters θ with respect to a new task while preventing catastrophic forgetting of previously learned tasks. GEM maintains a memory of past gradients and ensures that the current task's gradient does not deviate significantly from them. The optimization objective for GEM can be expressed as:

$$\theta_i^{**} = \arg\min_{\theta} L(f(\theta), D_i|\theta_{init} = \theta_{i-1}) + \lambda \sum cos\_sim(g_i, g_{i-1}) \tag{4.3}$$

  where, g_i represents the loss function's gradient for the i-th task with respect to θ. λ represents the hyperparameter which controls the regularization term's importance. $cos\_sim(g_i, g_{i-1})$ denotes the cosine similarity between the gradient of the current task g_i and the gradients of past tasks g_{i-1}.

  The objective function includes task-specific loss and a regularization term that penalizes changes in the gradient direction concerning past tasks. The regularization term encourages the current gradient to align with past gradients, thus preventing catastrophic forgetting while adapting to new tasks. The optimization seeks to minimize this combined objective to update the parameters θ for the current task in a way that balances learning the new task and retaining knowledge from previous tasks.

- **Synaptic Intelligence (SI)**: It is a technique used in continual learning to regularize the learning process by penalizing changes to important parameters

learned from previous tasks. It retains an approximate diagonal of the Fisher information matrix for each task and adjusts it as new tasks arise. The formula for SI involves calculating the importance of each parameter and applying regularization based on these importance scores. Here's the formula for SI:

$$\theta_i^{**} = \arg\min_{\theta} L(f(\theta), D_i|\theta_{init} = \theta_{i-1}) + \Omega(\theta_i, \theta_{i-1}) \tag{4.4}$$

$\Omega(\theta_i, \theta_{i-1})$ is the Regularization term based on synaptic importance, penalizing changes to important parameters learned from previous tasks while allowing flexibility for learning new tasks.

The regularization term Ω is calculated using the synaptic importance scores. It penalizes changes to important parameters learned from previous tasks while allowing flexibility for learning new tasks. The specific calculation of Ω involves maintaining a diagonal approximation of the Fisher Information matrix or other methods for estimating parameter importance scores.

- **Memory Aware Synapses (MAS)**: MAS is a method for preserving knowledge in neural networks by selectively updating parameters based on their importance for previous tasks. It assigns a memory importance score to each parameter and applies regularization proportional to these scores.

  The aim is to find the parameters θ_i^{**} that minimizes the loss on task i, while preserving knowledge learned from previous tasks by selectively updating parameters.

$$\theta_i^{**} = \arg\min_{\theta} L(f(\theta), D_i|\theta_{init} = \theta_{i-1}) + \Lambda \sum_{j=1}^{p} MAS(w_j, \theta_{i-1}) \tag{4.5}$$

  where Λ is the regularization strength, p is the total number of parameters, w_j denotes the importance score associated with parameter j, and $MAS(w_j, \theta_{i-1})$ is a function that computes the regularization term based on the importance score w_j and the parameters θ_{i-1} learned from the previous task.

The regularization term encourages the model to retain important parameters learned from previous tasks while updating parameters relevant to the current task. The importance scores w_j can be computed using various criteria, such as sensitivity to previous tasks or relevance to current and previous tasks.

4.2 Datasets

In our study aimed at evaluating the effectiveness of our proposed continual learning framework, we have meticulously selected three publicly available datasets that mirror the intricate dynamics of real-world financial payment systems. These datasets are particularly significant as they closely resemble the complexities commonly encountered in Enterprise Resource Planning (ERP) accounting data. We have undertaken thorough data preparation steps to ensure the suitability and reliability of these datasets for our analysis.

- The City of Philadelphia Payments (D_A): Dataset D_A comprises an extensive compilation of 238,894 payment records sourced from 58 distinct city departments operating within the City of Philadelphia. Each payment transaction in this dataset is meticulously documented with a comprehensive array of attributes. These attributes include 10 categorical variables capturing various transactional aspects such as payment type, departmental category, and vendor details, along with 1 numerical attribute providing insights into the monetary aspect of each transaction.

- The City of Chicago Payments (D_B): Dataset D_B encompasses a vast repository of 399,158 payment entries originating from 54 diverse city departments within the City of Chicago. Much like dataset DA, DB exhibits a rich blend of categorical and numerical attributes, which provide a detailed representation of each payment transaction. Specifically, it comprises 6 categorical attributes

offering insights into transactional characteristics and 1 numerical attribute providing quantitative information about the payments.

- The City of York Payments (D_C): Dataset D_C presents a detailed record of 202,026 payments executed by 49 city departments over the time span from 2019 to 2022 in the City of York. Similar to other above datasets, it encompasses a combination of categorical and numerical attributes, offering a holistic view of the payment ecosystem within the city. Notably, it features 6 categorical attributes elucidating transactional nuances and 2 numerical attributes providing additional quantitative dimensions to the payment data.

The datasets we've selected offer a rich tapestry of real-world financial transactions, reflecting the diverse and complex nature inherent in such systems. Within these datasets, we observe a spectrum of transaction types, ranging from individual manual payments to structured payment run records. [14] This diversity is crucial as it mirrors the heterogeneous nature of financial data encountered in enterprise environments, particularly in the realm of accounting and auditing processes.

Utilizing diverse transaction types, we aim to simulate realistic financial auditing scenarios in municipal organizations. Auditors scrutinize expenditures and payment runs for regulatory compliance and internal controls. By constructing simulation environments with varied datasets, we mimic real-world complexities, enhancing our continual learning framework's efficacy. Incorporating data from multiple municipalities like Philadelphia, Chicago, and York captures jurisdictional nuances while revealing common trends. This analysis improves auditing practices and risk mitigation strategies for policymakers and financial auditors in municipal settings.

4.3 Comprehensive Examination of Data

We commence our analysis by formally defining our data population, denoted as X, which comprises a collection of N journal entries [15]. Each entry, indexed from i =

1 to N, is represented as xi and encompasses a set of attributes. These attributes include both categorical (M attributes) and numerical (K attributes) components, capturing various details crucial to the journal entries, such as posting type, posting date, amount, and general ledger.

4.3.1 Identification of Anomalous Entries

We differentiate between two classes of anomalous entries[16]:

- **Global Anomalies:** Global anomalies are characterized by individual attribute values within journal entries that deviate significantly from typical patterns. These anomalies are relatively straightforward to detect due to their conspicuous nature and often indicate errors or irregularities in financial transactions. Here's a more detailed exploration:
  - Unusual Posting Times: Journal entries reflecting financial transactions posted at unconventional times, such as late at night or during weekends, are indicative of global anomalies. These deviations from regular business hours may signal potential errors in data entry or unauthorized activity.
  - Rarely Used Ledgers: Entries recorded in general ledger accounts that are seldom utilized for transactions stand out as global anomalies. Such occurrences may highlight irregular accounting practices, oversight, or even attempts to conceal fraudulent activities by using obscure ledger accounts.
  - Abnormal Transaction Amounts: Anomalously large or small transaction amounts compared to historical averages or established thresholds are indicative of global anomalies. These deviations may signify data entry errors, fraudulent activities such as embezzlement, or significant financial events requiring further investigation.

- **Local Anomalies:**Local anomalies involve subtle irregularities in the correlations among attribute values within journal entries, making them more challenging to detect than global anomalies. Perpetrators of fraudulent activities often attempt to mimic regular posting patterns, necessitating a nuanced approach to anomaly identification. Here's a deeper dive into local anomalies:
  - Unusual Co-occurrences: Journal entries featuring combinations of attribute values that rarely occur together represent local anomalies. For example, a specific document type posted in conjunction with atypical general ledger accounts and user accounts may indicate attempts to disguise fraudulent activities by mimicking legitimate transactions.
  - Inconsistent Posting Patterns: Anomalies in the frequency, timing, or sequence of transactions within journal entries suggest irregularities in posting patterns. Sudden fluctuations or deviations from established norms may signify attempts to manipulate financial data or conceal fraudulent activities.
  - Mismatched Transaction Details: Discrepancies between related attributes within journal entries point to local anomalies. For instance, inconsistencies between the posting type and associated transaction amount or discrepancies in general ledger categories may indicate data entry errors, deliberate manipulation, or fraudulent activities.

4.4 Preprocessing

$\{A_j\}_{j=1}^{M}$ represents the set of city departments from which payments originate. Here M represents the total number of city departments included in the dataset. Each department contributes to the dataset by generating payments, reflecting the diverse organizational structure and functional areas within the municipal setup. It encapsulates the entire spectrum of departments contributing to the dataset, providing a comprehensive view of the organizational landscape within the municipality.

It serves as a fundamental entity in our analysis, as it delineates the sources of payments and reflects the organizational diversity inherent in financial transaction data [16]. By examining payments originating from different city departments, we gain insights into the distribution, volume, and nature of financial activities across various functional areas.

Each A_j represents an individual city department within the dataset. These departments encompass various administrative units, such as finance, public works, utilities, and social services, among others. Each department is responsible for executing payments related to its respective functions and activities.

Furthermore, understanding the composition and characteristics of the departments generating payments enables us to tailor our analysis to specific organizational domains, identify trends or anomalies within individual departments, and assess the overall financial health and performance of the municipality.

For each of the three datasets, we undertake the subsequent preprocessing [16] procedures:

- **Identification of Top Departments (D Departments):** To initiate the data preparation process, we first identify the top D departments within each dataset. These departments are selected based on their transaction volumes, as indicated by the number of payments associated with each department. By focusing on departments with the highest payment volumes, we ensure that our analysis targets entities that significantly contribute to the overall financial activity within the respective municipalities. These departments, denoted as $\{A_k^{**}\}_{k=1}^{D}$, represent focal points for our analysis due to their significant transaction volumes.

- **Sampling Payments for Analysis (P Payments):** After identifying the top departments, we proceed to sample payments from each department for further analysis. The number of payments sampled from each department

is denoted as P. Through random sampling, we select a subset of payments from each department, capturing a representative sample of the transactional data. This sampling approach ensures that our analysis encompasses a diverse range of payment scenarios and transaction types. Subsequently, for each department A_k^{**} we perform a random sampling process to select M payments for further analysis. These sampled payments are represented as $x_i \in \{A_k^{**}\}_{k=1}^{D}$ where $i \in [1, P]$.

- **Transformation of Attribute Values:**
  - **Conversion of Categorical Attributes:** Categorical attributes within the sampled payments are transformed into a suitable numerical format to facilitate analysis and modeling. One-hot encoding is employed to convert categorical attribute values into binary representations. This transformation creates a series of binary features, with each feature corresponding to a unique attribute value within the categorical variable. By converting categorical attributes into a binary format, we enable machine learning algorithms to effectively interpret and analyze these attributes as input features.

    Categorical attribute values x_i^j within the sampled payments are transformed into one-hot numerical tuples of bits. This transformation, denoted as $\tilde{x}_i^j \in \{0, 1\}^\alpha$, ensures that each categorical attribute is represented in a binary format. In this context, α denotes the quantity of distinct attribute values found within the initial categorical attribute x^j.
  - **Conversion of Numerical Attributes:** Numerical attributes within the sampled payments undergo scaling to ensure uniformity and comparability across different attributes. Min-max scaling is employed to normalize numerical attribute values within a predetermined range, usually [0, 1]. This technique maintains the relative proportions between attribute values while ensuring they all lie within a uniform range. By scaling numerical attributes, it improves the stability and convergence

of machine learning algorithms, facilitating more efficient data modeling and analysis.

Numerical attribute values x_i^l are scaled to a standardized range using min-max scaling. The scaling formula

$$\tilde{x}_i^l = \left(\frac{x_i^l - \min\left(x^l\right)}{\max\left(x^l\right) - \min\left(x^l\right)} \right) \tag{4.6}$$

ensures that numerical attributes are uniformly distributed within the range of [0, 1]. Here, represent the minimum and maximum values, respectively, observed across all attribute values in x^l.

In all our experiments, we'll utilize 15 city departments (D) and 10,000 samples (P) to generate pre-processed journal entries ($\tilde{x}$).

To mimic a constantly changing setting, we structure the information into a sequence of N experience occurrences represented as $\{E_i\}_{i=1}^{N}$. By structuring the data into experiences, we mimic the temporal nature of financial transactions and capture the changing patterns and trends over time. Experiences can align with various timeframes, like financial quarters or years, mirroring the dynamic nature of financial data in real-world auditing scenarios.

Within each experience E_i, corresponding to dataset D_i, a subset of pre-processed payments is randomly sampled from the transformed data. For every department D_i, P/N pre-processed payments is drawn from each set of transactions $\tilde{x}_i \in \{A_k^{**}\}_{k=1}^{D}$, means from each department A_k^{**}.

In real-world audit setups, financial data is analyzed over distinct time periods to assess an organization's health and compliance. Organizing data into experiences aligns with audit practices, providing snapshots for anomaly detection. This temporal arrangement tracks financial behavior changes, identifies patterns, and detects anomalies indicating fraud or errors[16]. Structuring experiences and sampling payments ensures continual learning in anomaly detection. The model receives diverse

data, adapting to evolving patterns and improving real-time detection. This approach enhances anomaly detection's effectiveness and reliability in financial data.

4.5 Experimental Setup

The encoder network q_ψ and the decoder network p_ϕ in the Autoencoder Networks (AEN) utilize Leaky-ReLU activation functions with a scaling factor $\gamma = 0.4$. Activation functions introduce non-linearity into neural networks, enabling them to learn complex patterns in data.The Leaky-ReLU activation function is a modification of the Rectified Linear Unit (ReLU), which addresses the issue of "dying ReLU" where neurons become inactive and cease learning. It achieves this by introducing a small, non-zero gradient when the input is negative, preventing the gradient from becoming completely flat during backpropagation. This slight slope for negative inputs improves gradient flow, enabling better learning in neural networks by maintaining non-linearity and facilitating the detection of complex patterns in data.

In bottleneck layer of the decoder network, tanh activation functions are used. tanh (hyperbolic tangent) is another activation function commonly used in neural networks. It maps input values to the range [-1, 1], making it suitable for modeling data that is normalized or standardized. By constraining the output to this range, tanh ensures that reconstructed data falls within reasonable bounds, which can be particularly important for financial accounting data where certain attributes may have predefined ranges or boundaries.

The parameter sets ψ, ϕ of the AEN models are initialized following the recommendations by. This initialization scheme helps prevent vanishing or exploding gradients, common issues in deep learning. By initializing parameters in a way that ensures activations and gradients are neither too large nor too small, the model can learn more effectively and converge faster during training.

The model undergoes 500 epochs of training, with each batch containing 128 journal entries. Early stopping is implemented to prevent overfitting, halting training once the loss stabilizes. Optimization is achieved through Adam, a technique that merges the benefits of Adagrad and RMSprop, adjusting learning rates per parameter and adapting based on gradient moments. Adam's efficiency and efficacy in training deep learning models make it widely favored.

In all our experiments, we fixed the number of experiences to $N = 10$. During training, we calculate the binary cross-entropy error for each encoded journal entry. This error metric quantifies the difference between the original input $\tilde{x}_i$ and its reconstruction by the AEN model $\tilde{\tilde{x}}_i$. Here, BCE error is defined as,

$$L_{Rec}(\tilde{\tilde{x}}_i, \tilde{x}_i) = \frac{1}{N} \sum_{i=1}^{N} \tilde{\tilde{x}}_i \cdot \log(\tilde{x}_i) + \left(1 - \tilde{\tilde{x}}_i\right) \cdot \log(1 - \tilde{x}_i) \tag{4.7}$$

Minimizing this error encourages the model to learn representations that faithfully capture the salient features of the input data and generate accurate reconstructions.

To rigorously evaluate the effectiveness of continual learning methods in anomaly detection within financial auditing datasets, we will systematically inject 100 anomalies into each experience [16]. Leveraging the Faker library, renowned for its ability to generate synthetic data with realistic characteristics, we will simulate diverse anomalies representing real-world irregularities. These anomalies will span various attributes such as transaction amounts, posting times, and vendor details. By employing Faker's capabilities, we ensure the creation of anomalies that closely mimic actual fraudulent activities. Subsequently, we will compute average metric values, including precision, recall, and F1-score, across all experiences, thereby facilitating a comprehensive assessment of the continual learning methods' performance under varied conditions. This standardized approach enhances reproducibility and consistency in our evaluation process.

To identify irregularities in each occurrence, we'll establish a threshold using statistical metrics like the mean and standard deviation of the reconstruction errors. These metrics help to define the error distribution for typical data [17]. One common approach is to set the threshold as a multiple of the standard deviation (e.g., 3 standard deviations) above the mean reconstruction error i.e $threshold = meanError + 3 * stdError$. This ensures that anomalies, which typically result in higher reconstruction errors, are captured while minimizing false positives. We will periodically recompute the above on a rolling window of recent data to adapt the threshold to changing conditions. This adaptive thresholding can help the model adapt to concept drift or changes in the data distribution over time. Setting a threshold based on statistical measures allows you to define a range within which most normal data points are expected to fall. Anomalies, which typically result in higher reconstruction errors, can then be identified based on their deviation from this normal behavior.

4.6 Performance Metrics

4.6.1 Precision

Precision is a metric that gauges the accuracy of positive predictions made by a model. It assesses the proportion of correct positive predictions out of all the positive predictions made. In essence, precision answers the question: "Among the instances predicted as positive, how many are truly positive?" It's computed using the following formula:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{4.8}$$

A high precision indicates that the model has a low false positive rate, meaning it rarely misclassifies negative instances as positive.

4.6.2 Recall

Certainly, recall, also referred to as sensitivity or the true positive rate, assesses the ratio of correctly identified positive predictions to all actual positive instances in the dataset. It addresses the query: "What proportion of the actual positive instances did the model accurately identify?" Recall is computed using the formula:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{4.9}$$

A high recall indicates that the model effectively captures most positive instances without missing many.

4.6.3 F1 Score

The F1 score, a widely used metric in classification tasks, serves as a balanced measure of precision and recall. It proves especially valuable when dealing with datasets where the number of positive and negative instances is imbalanced. The F1 score is computed as the harmonic mean of precision and recall, ensuring equal importance to both metrics in the evaluation process.

$$\text{F1 Score} = 2 \cdot \left(\frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \right) \tag{4.10}$$

The F1 score ranges from 0 to 1, where 1 indicates perfect precision and recall, and 0 indicates poor performance in both metrics.

Chapter 5

Results and Discussion

Table 5.1: Avg. L_{Rec} For Each Experience

Datasets	CL Methods	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
DA	GEM	3.68	3.62	2.91	2.45	2.32	2.16	1.95	1.84	1.67	1.59
	SI	3.71	3.72	2.86	2.43	2.07	1.96	1.88	1.76	1.71	1.53
	MAS	3.65	3.19	2.43	2.1	1.92	1.81	1.75	1.68	1.63	1.47
DB	GEM	3.26	3.18	3.02	2.72	2.12	1.83	1.74	1.52	1.43	1.39
	SI	3.35	3.25	3.16	2.64	2.13	1.77	1.61	1.49	1.46	1.43
	MAS	3.13	3.09	2.89	2.57	1.98	1.65	1.54	1.41	1.37	1.29
DC	GEM	3.64	3.53	3.42	3.37	2.92	2.34	2.12	1.98	1.85	1.77
	SI	3.67	3.62	3.59	3.29	2.98	2.41	2.19	1.91	1.83	1.72
	MAS	3.59	3.22	3.03	2.99	2.71	2.31	1.98	1.83	1.75	1.61

The provided reconstruction loss data across ten experiences as per Table 5.1 for GEM, SI, and MAS continual learning methods reveals promising trends in anomaly detection on financial dynamic datasets. Overall, all methods exhibit decreasing reconstruction losses over experiences, indicating effective adaptation to evolving data distributions. GEM consistently demonstrates the lowest losses, suggesting robust performance and efficient anomaly detection. SI and MAS also exhibit competitive performance, with decreasing losses over time. Variability in losses among experiences highlights fluctuations in dataset complexity and distribution. Despite higher losses in earlier experiences, all methods showcase improved performance in later experiences, indicating adaptability and stability in capturing normal data patterns

and detecting anomalies. Visual representation of this can be found in Figure 5.1 and Figure 5.2. These findings underscore the potential of continual learning methods for enhancing fraud detection and prevention in financial auditing, offering valuable insights for real-world applications.

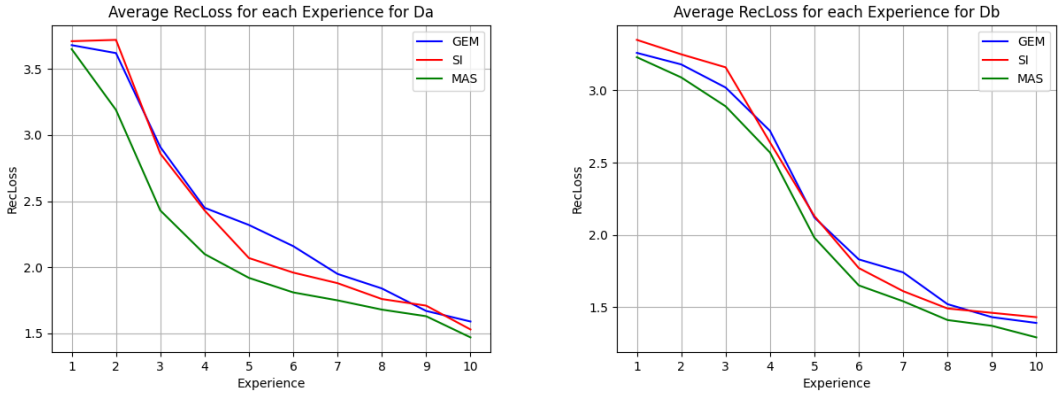

FIGURE 5.1: Average Reconstruction Loss For Each Experience in DA and DB

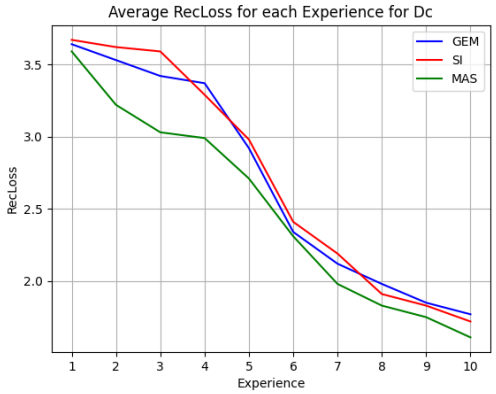

FIGURE 5.2: Average Reconstruction Loss For Each Experience in DC

Furthermore, we have also calculated the overall average of anomaly detection precision and recall across all experiences for the injected anomalies in each scenario. This additional evaluation step provides insights into the collective performance of the continual learning methods in detecting anomalies across multiple experimental conditions. By aggregating precision and recall metrics, we gain a comprehensive understanding of the methods' ability to accurately identify anomalies while minimizing false positives and negatives. This holistic analysis strengthens the validity and reliability of our assessment, offering valuable insights into the robustness of

the continual learning approaches in anomaly detection within financial auditing datasets.

Table 5.2: Evaluation Metrics for the Injected Anomalies

Datasets	Methods	Avg. Precision	Avg. Recall	F1 Score
DA	**GEM**	54.9	64.9	59.5
	SI	56.2	68.1	61.6
	MAS	61.7	73.3	67.0
	Baseline [16]	44.5	50.2	47.2
DB	**GEM**	65.9	86.3	74.7
	SI	67.6	87.2	76.1
	MAS	73.4	90.8	81.2
	Baseline [16]	56.6	69.4	62.4
DC	**GEM**	58.5	73.6	65.2
	SI	57.7	72.7	64.3
	MAS	64.8	79.9	71.6
	Baseline [16]	42.9	63.4	51.2

Table 5.2 showcases the performance analysis of three continual learning (CL) methods – GEM, SI, and MAS – across three distinct financial auditing datasets (DA, DB, DC). The metrics, including average precision, recall, and F1 score, offer valuable insights into the effectiveness of each CL method in detecting injected anomalies within each experience.

MAS consistently demonstrates the highest average precision and recall across all datasets, indicating its superior performance in accurately identifying anomalies while minimizing false positives and false negatives. This suggests that MAS is adept at distinguishing between normal and abnormal transactions within financial auditing data. Its robust performance underscores its potential for real-world applications in anomaly detection tasks.

The findings emphasize the importance of selecting the appropriate CL method for effective anomaly detection in financial auditing datasets. MAS emerges as a promising choice due to its consistently high precision and recall values. Leveraging

MAS can enhance the accuracy and reliability of anomaly detection processes, ensuring the timely identification of fraudulent activities and safeguarding organizational resources.

Chapter 6

Conclusion and Future Scope

This thesis has demonstrated the effectiveness of employing Adversarial Exemplar Networks (AENs) with continual learning (CL) methods like GEM, SI, and MAS for anomaly detection in dynamic financial data streams. Through rigorous experimentation, it has been established that this approach significantly enhances performance while addressing challenges faced by traditional machine learning techniques. AENs coupled with CL enable models to adapt to evolving data distributions and concept drift, thereby improving detection accuracy. Real-time adaptation mechanisms further enhance the models' ability to discern anomalies in dynamic environments. This research underscores the importance of innovative methodologies in anomaly detection and highlights the potential of AENs with CL for enhancing anomaly detection systems' reliability in financial applications. Future research directions may explore additional CL methods and validate the proposed approach in real-world financial scenarios.

Future research could explore integrating federated learning for collaborative model training across distributed data sources while preserving privacy. Reinforcement learning algorithms could enhance adaptability by optimizing detection strategies over time. Advancements in deep generative models like VAEs and GANs offer

potential for capturing complex data distributions and improving anomaly detection. Hybrid approaches combining multiple techniques, such as ensemble methods or hybrid neural architectures, may further enhance detection performance. Additionally, investigating advanced anomaly interpretation techniques like attention mechanisms or explainable AI methods could increase model transparency and human understanding of detected anomalies, boosting trust in financial applications.

Bibliography

[1] Y. Song, S. Hyun, and Y.-G. Cheong, "Analysis of autoencoders for network intrusion detection," *Sensors*, vol. 21, no. 13, p. 4294, 2021.

[2] L. Xu, X. Ding, H. Peng, D. Zhao, and X. Li, "Adtcd: An adaptive anomaly detection approach towards concept-drift in iot," *IEEE Internet of Things Journal*, 2023.

[3] G. Li and J. J. Jung, "Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges," *Information Fusion*, vol. 91, pp. 93–102, 2023.

[4] D. Samariya and A. Thakkar, "A comprehensive survey of anomaly detection algorithms," *Annals of Data Science*, vol. 10, no. 3, pp. 829–850, 2023.

[5] Q. Xiang, L. Zi, X. Cong, and Y. Wang, "Concept drift adaptation methods under the deep learning framework: A literature review," *Applied Sciences*, vol. 13, no. 11, p. 6515, 2023.

[6] A. Seetha, S. S. Chouhan, E. S. Pilli, and V. Raychoudhury, "D i e v d: Disruptive event detection from dynamic datastreams using continual machine learning: A case study with twitter," *IEEE Transactions on Emerging Topics in Computing*, 2023.

[7] A. Lee, Y. Zhang, H. M. Gomes, A. Bifet, and B. Pfahringer, "Look at me, no replay! surprisenet: Anomaly detection inspired class incremental learning,"

in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 2023, pp. 4038–4042.

[8] D.-W. Zhou, H.-L. Sun, J. Ning, H.-J. Ye, and D.-C. Zhan, "Continual learning with pre-trained models: A survey," *arXiv preprint arXiv:2401.16386*, 2024.

[9] L. Wang, X. Zhang, H. Su, and J. Zhu, "A comprehensive survey of continual learning: Theory, method and application," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.

[10] M. Mundt, Y. Hong, I. Pliushch, and V. Ramesh, "A wholistic view of continual learning with deep neural networks: Forgotten lessons and the bridge to active and open world learning," *Neural Networks*, vol. 160, pp. 306–336, 2023.

[11] P. Li, Y. Pei, and J. Li, "A comprehensive survey on design and application of autoencoder in deep learning," *Applied Soft Computing*, vol. 138, p. 110176, 2023.

[12] K. Berahmand, F. Daneshfar, E. S. Salehi, Y. Li, and Y. Xu, "Autoencoders and their applications in machine learning: a survey," *Artificial Intelligence Review*, vol. 57, no. 2, p. 28, 2024.

[13] J. Yu, X. Gao, F. Zhai, B. Li, B. Xue, S. Fu, L. Chen, and Z. Meng, "An adversarial contrastive autoencoder for robust multivariate time series anomaly detection," *Expert Systems with Applications*, vol. 245, p. 123010, 2024.

[14] Z. Sun, J. Wang, and Y. Li, "Ramfae: a novel unsupervised visual anomaly detection method based on autoencoder," *International Journal of Machine Learning and Cybernetics*, vol. 15, no. 2, pp. 355–369, 2024.

[15] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: a review of anomaly detection techniques and recent advances," *Expert systems With applications*, vol. 193, p. 116429, 2022.

[16] M. Schreyer, T. Sattarov, D. Borth, A. Dengel, and B. Reimer, "Detection of anomalies in large scale accounting data using deep autoencoder networks," *arXiv preprint arXiv:1709.05254*, 2017.

[17] R. Lehmann, "3 σ-rule for outlier detection from the viewpoint of geodetic adjustment," *Journal of Surveying Engineering*, vol. 139, no. 4, pp. 157–165, 2013.