Deception Detection: From Static Texts to Multimodal Signals

Mandela Logan

mandelakorilogan@gmail.com

Abstract

Fraud detection remains a critical challenge in various domains, including finance, telecommunications, and e-commerce, where malicious actors employ sophisticated methods to evade detection. This paper provides a comprehensive reflection on techniques applied in two fraud detection projects: a static text-based model and a multimodal audio-text model. Drawing from practical implementations, we delve into data preprocessing, feature engineering, modelling, calibration, validation, and explainability for each approach. Shared engineering practices, such as modular repository structures and experiment tracking, are also discussed. By integrating classical statistical methods with modern deep learning paradigms, these techniques underscore the balance between efficiency, interpretability, and robustness in real-world fraud detection systems. Insights from these projects highlight pathways for advancing applied AI in high-stakes security applications.

1 Introduction

In an era of escalating digital transactions, fraud poses significant economic and reputational risks to organizations. Traditional rule-based systems, while reliable, often fail against adaptive fraudsters who obfuscate patterns through text manipulations or vocal disguises. This paper reflects on techniques employed in two distinct fraud detection projects: one focused on static text analysis (e.g., emails or messages) and another on multimodal analysis incorporating audio and transcribed text (e.g., call center interactions). The static approach emphasizes efficiency and interpretability for high-volume screening, while the multimodal model leverages richer behavioural cues for nuanced detection.

The reflections are structured as follows: Section 2 explores the evolving nature of digital deception, Section 3 delves into static fraud detection in depth, Section 4 covers multimodal techniques, Section 5 discusses shared engineering practices across both projects, and Section 6 offers a final perspective on their implications. These insights are derived from iterative project developments, aiming to bridge theoretical AI advancements with practical deployment challenges.

2 The Evolving Fraud Landscape: The New Normal in Digital Deception

The digital threat landscape is evolving at a rapid pace, with cybercriminals harnessing cutting-edge technologies such as artificial intelligence, deepfakes, and advanced social engineering to create more complex and challenging fraud tactics.[1] The analysis of fraud must begin by understanding the nature of this modern adversary.

2.1 The Rise of AI-Enabled Fraud

The current state of AI-driven fraud is a significant departure from previous eras, where simple, manually updated rule sets were sufficient. The threat has shifted from static, predictable methods to dynamic, AI-enabled schemes. The use of generative AI (GenAI) and large language models (LLMs) allows malicious actors to scale their operations with minimal human oversight. For example, AI-generated emails can now mimic legitimate communication with remarkable accuracy, eliminating the grammatical errors and awkward phrasing that often revealed phishing attempts in the past. The fraud landscape has seen a staggering increase in AI-enabled scams, with reports indicating a rise of over 450% between May 2024 and April 2025.[1] A particularly dangerous trend is the use of deepfakes and biometric spoofing. Highly convincing deepfake videos and voice samples can impersonate trusted individuals, enabling sophisticated social engineering attacks. A notable case in Hong Kong involved a company being defrauded of $25 million by deepfake impersonations of executives.[1] This form of deception necessitates a move beyond traditional identity verification methods that rely on physical characteristics alone. The traditional security paradigm, often described as a "fortress defence," relied on fixed rules and a static blacklist of known fraudulent behaviours.[1] The assumption was

that by identifying the characteristics of past fraud, one could pre-emptively block future attempts. However, as fraudsters learned these rules, they adapted their tactics to circumvent them. For instance, a rule that flags any transaction over a certain amount could be easily bypassed by a fraudster splitting a single large transaction into multiple smaller ones that fall below the threshold.[4] The emergence of GenAI has accelerated this adversarial feedback loop. A static, rule-based defence is fundamentally mismatched against a generative, constantly evolving attack vector. The new challenge is not just detecting known bad patterns but identifying a vast, unknown, and continuously morphing set of anomalies. This requires a shift from a reactive, manually updated blacklist to a self-learning system that can autonomously adjust to new threats.[2]

2.2 A Shift from Bots to Behavioural Analysis

The escalating sophistication of fraudulent activities has necessitated a profound change in the fraud detection paradigm. The focus is no longer a simple binary classification of "human" versus "bot" but rather a nuanced analysis of "legitimate intent" versus "fraudulent intent".[2] This behavioural approach evaluates user actions in real time against established profiles of expected behaviour, flagging any significant deviations as suspicious.[4] An effective, modern fraud detection strategy requires a multi-layered approach that goes beyond basic single-point checks. It combines real-time monitoring and advanced behavioural analytics with other security measures, such as multi-factor authentication (MFA), risk-based authentication (RBA), and device fingerprinting.[2] For example, a system might flag a transaction based on an anomalous location but then use device fingerprinting or behavioural biometrics to confirm the user's identity. This layered defence strengthens the overall system's resilience and reduces false positives, which are a major source of customer friction. The fundamental change in the adversarial landscape has made the development of a reactive, rule-based defence an unsustainable strategy. The challenge now is to build systems that not only detect fraud but also understand the behavioural dynamics and intent behind it, justifying the need for the advanced, full-stack approaches explored in the remainder of this report.

3 Foundational Techniques in Text-Based Fraud Detection

Static fraud detection relies on analyzing textual content without temporal or auditory context, making it suitable for scalable, low-latency applications like email filtering or transaction monitoring. The process is a full-stack endeavour, encompassing everything from foundational data cleaning to sophisticated modelling and validation.

3.1 Advanced Data Preprocessing and Obfuscation

Effective preprocessing is foundational to handling noisy, unstructured text data common in fraud scenarios. In the context of fraud, this is not a simple data cleaning task; it is an adversarial process designed to reverse the obfuscation tactics employed by fraudsters.[8][9][19][20]

3.1.1

Text normalization involves a series of steps to standardize text. This includes converting all characters to lowercase and applying Unicode normalization (e.g., NFKC) to standardize accented or variant characters, ensuring consistency across diverse inputs.[9] Whitespace standardization collapses multiple spaces into single ones, reducing variability.[9] Irrelevant elements are removed or normalized: HTML tags are stripped, emojis are either removed or mapped to semantic tokens (e.g., ":smile:"), and non-printable characters are filtered to prevent model confusion.[9]

3.1.2 The Leetspeak and Euphemism Problem

A key challenge in text-based fraud detection is the use of obfuscated language by fraudsters to evade keyword-based filters. Leetspeak (1337), for example, replaces letters with numbers or symbols that resemble them (e.g., "p@ssw0rd").[10][11] Programmatic normalization for this requires a systematic approach, often involving a dictionary-like mapping of common substitutions. A simple script can iterate through a string, replacing characters like 'a' with '4' or 'e' with '3'.[12] More advanced solutions, such as the pyleetspeak library, offer fine-grained control over replacement probabilities and the ability to handle more complex substitutions, even for entire words.[13]

Beyond leetspeak, fraudsters also use euphemisms to bypass content filters, such as using "unalive" for "kill" or "seggs" for "sex". Detecting these requires a more sophisticated, context-aware approach, such as using

semantic embeddings or a regularly updated lexicon of known euphemisms. The effectiveness of a fraud detection model hinges on its ability to handle this continuous evolution of deceptive language.[20]

3.1.3 Sanitization of Sensitive Information

URL, email, and phone number masking is a critical preprocessing step for two primary reasons: privacy and model generalization. Replacing these sensitive entities with non-sensitive placeholders (e.g., ``,[PHONE]) prevents the model from learning patterns that are tied to specific, private data.[14][15][16] The model is instead forced to learn the pattern of a fraudulent message, such as the presence of multiple URLs in a short message, a common tactic in phishing scams, without leaking sensitive information or being overfit to a specific instance. This process can be implemented using regular expressions or specialized NLP libraries.[17][18]

3.2 Feature Engineering: Beyond Term Frequency

Feature engineering transforms raw text into discriminative representations that can be used by machine learning models.

3.2.1 Classic Approaches (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) is a powerful, classic technique that weighs the importance of a term in a document relative to a corpus of documents. In fraud detection, it is applied at both the character n-gram and word n-gram levels. Character n-grams (e.g., 3-6 grams) are particularly effective at detecting leetspeak and other forms of character obfuscation. Word n-grams (e.g., 1-2 grams) capture semantic context and can flag suspicious phrases like "kindly transfer" or "urgent refund". TF-IDF's inverse document frequency component helps to downplay the significance of common terms (like "the" or "and") while highlighting rare, domain-specific words that are more likely to be indicators of fraud.[3][21]

3.2.2 Linguistic and Behavioural Features

NLP enables the extraction of deeper, more nuanced features from text. Lexicon features incorporate domain-specific dictionaries to flag high-risk words such as "refund," "claim," or "urgent." Message length features, including total character count, word count, and average word length, can capture the verbosity typical of persuasive fraud. Special character ratios (e.g., the proportion of digits or symbols) often spike in fraudulent messages, such as those with excessive exclamation marks. Beyond simple counts, advanced linguistic analysis can detect subtle cues. Sentiment and linguistic pattern recognition can flag unusual emotional distancing, excessive detail, or the use of a passive voice, all of which may indicate deceptive intent. Entity and relationship mapping can also be used to uncover coordinated fraud networks by identifying relationships between seemingly unrelated claimants or entities.[21]

3.3 A Comparative Analysis of Modelling Paradigms

The choice of modelling technique for text-based fraud detection involves a strategic trade-off between performance, efficiency, and interpretability. The optimal solution is not about finding a single "best" model but about building a system that balances these factors to meet diverse operational and compliance needs.

3.3.1 Machine Learning

Traditional ML models, such as Logistic Regression, Naive Bayes, Support Vector Machines, and tree ensembles (Random Forest, LightGBM, XGBoost), have a long and successful history in fraud detection. They are particularly effective when applied to well-engineered, structured features like TF-IDF vectors or feature counts. These models are highly prized for their interpretability. For example, the coefficients of a Logistic Regression model can be inspected to reveal the weighted fraud indicators, which provide a clear and defensible reason for a decision. Similarly, tree-based models like XGBoost can provide feature importance scores. These models offer a good balance between accuracy and computational efficiency, making them suitable for real-time applications where low latency is critical.[3][5][22][24]

3.3.2 Deep Learning (DL) with Transformers

The advent of deep learning and, specifically, transformer-based models like BERT and DistilBERT has revolutionized text classification. These models excel at capturing complex sequential and contextual patterns in text, which can lead to higher accuracy on nuanced fraud. A BERT-based model can understand the semantic meaning of a message beyond simple keywords, such as detecting subtle insincerity or deception. However, these models are computationally expensive and often act as "black boxes," making their decision-making process difficult to interpret.[25][26][27]

3.3.3 The Hybrid Model as an Optimal Compromise

In high-stakes, regulated environments like finance, a model's ability to explain its decision is often as important as its accuracy. This has led to the rise of hybrid models that combine the strengths of both paradigms. A powerful approach involves using a transformer model (e.g., BERT) as a feature extractor to generate deep contextual embeddings, which are then fed into a classical, interpretable classifier like XGBoost. This strategy leverages the power of deep learning for complex pattern recognition while maintaining the explainability and computational efficiency of a tree-based model. It is a pragmatic solution that acknowledges that the most technologically advanced model is not always the most effective or appropriate for a given business context.[23]

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	Interpretability	Inference Latency
Logistic Regression	88.5	87.1	89.2	88.1	High	Low
XGBoost	93.8	92.5	94.1	93.3	Moderate	Low
BERT	97.1	96.2	97.5	96.8	Low	High
BERT-XGBoost Hybrid	98.4	97.9	98.6	98.2	Moderate	Moderate
Note: Metrics are representative and may vary significantly based on the dataset and task.[22][23][24]						

Table 1: Performance Benchmarks for Text Classification Models

4 Multimodal Fraud Detection: A Layered Defence with Audio and Text

Multimodal fraud detection integrates audio and text from sources like phone calls, capturing vocal nuances (e.g., hesitation) alongside textual content for superior accuracy. This approach provides a richer, more robust understanding of an interaction, as different data types can compensate for the weaknesses of others.[30][31]

4.1 The Value of Multimodality and the "Modality Gap"

Combining data modalities, such as analyzing a transaction amount alongside a user's geolocation and device fingerprint, can reveal inconsistencies that single-data approaches miss. In the context of call center interactions, a multimodal system can cross-validate a suspicious phrase in a transcript with a speaker's vocal cues.[30]

However, a significant challenge, known as the "modality gap," exists between raw audio and text-based deep learning models. While automatic speech recognition (ASR) can transcribe audio into text, this process can result in information loss, omitting crucial fraud indicators contained in vocal features like tone and pauses. This necessitates a full-stack approach that extracts and analyzes both types of data.[30]

4.2 Deep Audio Feature Extraction

Analyzing audio requires a specialized feature engineering pipeline that transforms raw audio signals into meaningful representations.

4.2.1 Classic Features (MFCCs and Prosodic Markers)

Mel Frequency Cepstral Coefficients (MFCCs) are a classic set of features widely used to represent the timbre of a voice. The calculation of MFCCs involves several steps:

1. Framing the Signal: The audio is divided into short, overlapping frames (e.g., 20-40 ms).
2. Fourier Transform: For each frame, a power spectrum is calculated using a Fourier transform.
3. Mel Filterbank: The power spectrum is passed through a Mel filterbank, which is a series of triangular filters that mimic the non-linear human auditory perception of pitch.
4. Logarithm and DCT: The logarithm of the filterbank energies is taken, and a Discrete Cosine Transform (DCT) is applied to decorrelate the features, yielding the MFCCs.[35][36]
   Beyond timbre, prosodic features capture the rhythm and intonation of speech. These include pitch (F0), jitter (cycle-to-cycle pitch variation), and shimmer (amplitude variation). These features are particularly valuable in fraud detection as they can reveal stress, deception, or unusual emotional states that may be missed in a simple text transcript.[37][38]

4.2.2 Learned Embeddings from Self-Supervised Learning

A breakthrough has been the use of self-supervised learning (SSL) models like wav2vec 2.0. These models learn rich, contextual representations directly from vast amounts of unlabelled audio data, overcoming the scarcity of labelled fraud data, which is a significant challenge due to privacy concerns. The model is pretrained on a large corpus of unlabelled speech and then fine-tuned on a smaller, task-specific dataset. This approach has been shown to be effective for anti-spoofing and other speech tasks, making it a powerful tool for fraud detection.[39][40][41][42]

4.3 Multimodal Fusion Strategies and Implementation

The success of a multimodal system hinges on how the different data streams are combined. The report analyzes three primary fusion strategies, each with its own trade-offs.[32][43][44][45][46][47][48]

Early Fusion

Early fusion concatenates raw or low-level features from all modalities before they are fed into a single model. For example, a concatenated vector of audio MFCCs and text TF-IDF features would serve as the input to a single classifier. This approach is simple to implement and can be computationally efficient. However, it is not well-suited for disparate modalities like audio and images, as there is no practical way to merge them into a single coherent tensor. It also risks the "curse of dimensionality" and may fail to capture complex, non-linear relationships between the modalities.

Late Fusion

In late fusion, each modality is processed independently by its own specialized model. The final decision is then made by combining the outputs of these individual classifiers, often through weighted averaging, stacking, or a simple voting mechanism. This strategy is robust to missing data; a model can still make a prediction even if one modality is absent, and offers greater interpretability, as the contribution of each modality to the final decision can be analyzed.

Mid-Fusion with Cross-Attention

Mid-fusion, particularly with cross-attention mechanisms, represents a state-of-the-art approach that seeks to overcome the limitations of early and late fusion. This strategy fuses features at an intermediate layer of the model, allowing for dynamic, learned interactions between modalities. A cross-attention mechanism uses queries from one modality (e.g., text) to "attend to" or find the most relevant information in another modality (e.g., audio). For example, the model could learn to correlate a keyword like "urgent" with a specific spike in the speaker's pitch. This provides a richer understanding of the interaction and allows the model to prioritize key segments for fraud detection. The development of multimodal systems is a direct response to a fundamental "data paradox" in high-stakes, private domains. While multimodal models are demonstrably superior, obtaining vast quantities of high-quality, labelled data that combines sensitive information like audio and text is prohibitively difficult due to privacy concerns and the cost of manual annotation. A groundbreaking approach to this challenge is the use of synthetic data generation. Datasets like TeleAntiFraud-28k use a multi-agent adversarial framework to simulate realistic fraud scenarios, generating large-scale, privacy-preserved, and annotated data.[33][34][49][50] This represents a shift from purely reactive data collection to proactive, model-driven data synthesis, ensuring that models can be trained on a diverse range of adversarial scenarios that they would otherwise never encounter. This fundamentally changes the nature of the training pipeline in critical security applications.

Strategy	Description	Pros	Cons
Early Fusion	Concatenates features or raw data from all modalities before feeding them into a single model.	Simple to implement; Computationally efficient.	May not capture complex inter-modal relationships; Susceptible to the curse of dimensionality.
Mid-Fusion (e.g., Cross-Attention)	Uses attention mechanisms to dynamically learn relationships between modalities at an intermediate layer of the model.	Captures deep, synergistic information between modalities; More flexible than early fusion.	Computationally intensive; More complex to implement.
Late Fusion	Processes each modality independently with its own classifier; it combines the final outputs for a single decision.	Robust to missing data in one or more modalities; More interpretable as each model's contribution is explicit.	Does not model the direct interaction between modalities.

Table 2: Multimodal Fusion Strategies

5 MLOps and Production: Bridging the Research-to-Reality Gap

A fraud detection model is not a static artifact; it is a dynamic system that must be continuously monitored and updated to remain effective. The MLOps pipeline is not a technical afterthought but a core business strategy for ensuring long-term value and competitive advantage.

5.1 The Criticality of Experiment and Artifact Versioning

In a fraud detection arms race, a model's effectiveness begins to degrade the moment it is deployed. This makes a single, one-off training effort insufficient. A robust MLOps pipeline transforms fraud detection from a one-time project into a scalable, maintainable, and continuously improving business capability.[53] Experiment tracking is a cornerstone of this process. It involves meticulously logging all relevant metadata for every experiment, including hyperparameters, model architectures, training procedures, and evaluation metrics. This ensures that every result is reproducible, allowing data scientists to understand the cause and effect of their changes and steer the development process in the right direction. Tools like MLflow or Weights & Biases are essential for this task, offering centralized, collaborative platforms to manage the iterative model development process.[54]

5.2 Continuous Monitoring and Drift Mitigation

The adversarial nature of fraud ensures that a model's performance will inevitably decay over time. This phenomenon is known as "model drift," which occurs when the data a model encounters in production deviates from the data it was trained on.[52] There are two primary types of drift:

- Data Drift (Covariate Shift): This occurs when the distribution of the input features changes over time. For example, fraudsters might start using a new vocabulary or a different set of obfuscation techniques, shifting the input distribution away from what the model learned during training.[51]
- Concept Drift: This is a more severe form of drift where the relationship between the input features and the target variable changes. For example, a legitimate transaction pattern could become a fraudulent one over time due to a new fraud scheme.[51][52]F

Proactive monitoring is critical for detecting and mitigating drift before it leads to significant business losses. This can be achieved through automated monitoring dashboards that track key performance metrics (e.g., accuracy, precision, recall) and statistical indicators of drift (e.g., KL divergence). When a model's performance drops below a predefined threshold, an alert can be triggered to prompt human review or, in a fully automated system, to initiate a retraining process. Mitigation strategies include scheduled retraining, which updates the model regularly, and triggered retraining, which automatically initiates a new training run when drift is detected. Advanced approaches also involve using synthetic data to test a model's robustness against new, hypothetical fraud scenarios before they appear in the wild.

5.3 Real-World Case Studies in Fraud Prevention

The value of a robust, AI-driven fraud detection strategy is demonstrated by tangible business outcomes. By combining advanced models with a rigorous MLOps pipeline, organizations can achieve significant fraud reduction and operational efficiency. One case study from a UK retail bank saw a solution that integrated machine learning and AIOps to analyze unstructured data from various sources. The results included a reduction in fraud incidents, enhanced security visibility, and optimized investigation workflows through automation.

Furthermore, these systems can significantly reduce false positives, which improves the customer experience and saves costs. A single-model approach can be more vulnerable to degradation, as a successful fraudster tactic can diminish the value of the entire model. In contrast, a "model-of-models" or a multiple-model approach is more robust because the failure of one model does not degrade the performance of the others. A case study for a financial institution showed that a multiple-model approach provided a 3.3% increase in fraud value detection, which translated to nearly $1 million

in additional fraud detected annually. Another example of a successful production deployment reported a 10% increase in checkout success rates due to a decline in false positives, which directly improved customer satisfaction.[53]

In fraud detection, the MLOps pipeline is not just a technical detail; it is the core business strategy for ensuring long-term value and competitive advantage. The true measure of a fraud detection system's success is not its initial accuracy score but its ability to sustain that performance in a dynamic, hostile, and constantly evolving production environment.

6 Conclusion

By combining classic statistical methods (TF-IDF, Logistic Regression, MFCCs) with modern deep learning paradigms (transformers, wav2vec, attention fusion), the two projects discussed in this paper illustrate a comprehensive, full-stack approach to fraud detection. The static model prioritizes efficiency, interpretability, and rapid deployment, making it ideal for high-volume, text-heavy environments. The multimodal model, on the other hand, adds depth, robustness, and behavioural insight by integrating vocal cues with textual content, excelling in interactive scenarios.[28][29]

6.1 A Full-Stack Approach: The Power of Integration

The most robust fraud detection systems are not built on a single, "best" model but on a hybrid architecture that leverages the strengths of multiple approaches. They use traditional ML on structured features for efficiency and explainability while reserving the power of deep learning for complex, unstructured, and multimodal data. This integrated strategy, coupled with a rigorous MLOps pipeline, ensures that a system can not only detect sophisticated fraud but also maintain its effectiveness over time in the face of an adaptive adversary.

6.2 Future of Fraud Detection

The battle against fraud is a continuous arms race. The future of fraud detection will be defined by a relentless focus on adaptability and proactive defence. The maturation of Large Audio Language Models (LALMs) that can directly process audio signals without a lossy ASR step will bridge the "modality gap" and enable more accurate and nuanced detection of fraud. Furthermore, the development of sophisticated data synthesis methods, such as agent-based adversarial simulation, will be crucial for creating richer, more realistic training data that anticipates future fraud tactics, ensuring that models are trained on what's coming, not just on what has already happened.[28]

Together, these techniques position the system not only as practical but also at the frontier of applied AI for fraud detection, paving the way for hybrid systems that can truly adapt to emerging threats.

Aspect	Static Model	Multimodal Model
Use Case	E-commerce, messages, emails	Call centers, interactive systems
Data Modalities	Text	Audio, Text
Primary Objective	Efficiency and speed	Depth and nuanced insight

Aspect	Static Model	Multimodal Model
Key Techniques	TF-IDF, N-grams, XGBoost, Transformers	MFCCs, Wav2vec, Attention Fusion
Efficiency	High (low latency, low compute)	Moderate (higher latency, requires GPUs)
Interpretability	High (e.g., feature importance, coefficients)	Moderate (e.g., attention weights, prosodic visualizations)
Robustness	Moderate (vulnerable to subtle obfuscation)	High (can cross-validate across modalities)

Table 3: Comparative Analysis of Static vs. Multimodal Fraud Detection

Acknowledgements

I would like to express my gratitude to Umberto Mautone for his inspiring work as one of the engineers behind the SmartShield™, by Veriswitch Solutions Inc., which served as a significant source of inspiration for this project. I also thank Concillia Muonde for her valuable insights and guidance on intellectual property matters. Additionally, I am grateful to Umberto Mautone for providing the A100 GPU essential for the training, testing, and research efforts of this work.

Works Cited

[1] Digital Deception: Fighting Fraud in the Era of Emerging Technology | Alvarez & Marsal, https://www.alvarezandmarsal.com/thought-leadership/digital-deception-fighting-fraud-in-the-era-of-emerging-technology
[2] How AI Is Used in Fraud Detection in 2025 - DataDome, https://datadome.co/learning-center/ai-fraud-detection/
[3] Machine Learning for Fraud Detection: An In-Depth Overview - Itransition, https://www.itransition.com/machine-learning/fraud-detection
[4] What Are Fraud Detection Rules: Types and Best Practices - Vespia, https://vespia.io/blog/fraud-detection-rules
[5] How machine learning works for payment fraud detection and prevention - Stripe, https://stripe.com/resources/more/how-machine-learning-works-for-payment-fraud-detection-and-prevention
[6] Credit Card Fraud Detection Case Study | SPD Technology, https://spd.tech/machine-learning/credit-card-fraud-detection-case-study/
[7] Detecting Scams Using Large Language Models - arXiv, https://arxiv.org/html/2402.03147v1
[8] Text Preprocessing in NLP with Python Codes - Analytics Vidhya, https://www.analyticsvidhya.com/blog/2021/06/text-preprocessing-in-nlp-with-python-codes/
[9] NLP — Text PreProcessing (Part 1) | by Chandu Aki | The Deep Hub | Medium, https://medium.com/thedeephub/nlp-text-preprocessing-part-1-dfc7d3ee0977
[10] Leet - Wikipedia, https://en.wikipedia.org/wiki/Leet
[11] What is leetspeak? - A Beginner's Guide | Lenovo US, https://www.lenovo.com/us/en/glossary/what-is-a-leetspeak/

[12] Personal-Python-Scripts/leetSpeak.py at master - GitHub, https://github.com/CuriousLearner/Personal-Python-Scripts/blob/master/leetSpeak.py
[13] pyleetspeak · PyPI, https://pypi.org/project/pyleetspeak/
[14] Masking private user information using Natural Language Processing - ResearchGate, https://www.researchgate.net/publication/352786519_Masking_private_user_information_using_Natural_Language_Processing
[15] What Is Tokenization? - Akamai, https://www.akamai.com/glossary/what-is-tokenization
[16] What is Tokenization | Data & Payment Tokenization Explained - Imperva, https://www.imperva.com/learn/data-security/tokenization/
[17] 5 Best Python NLP Libraries in 2025 - Kommunicate, https://www.kommunicate.io/blog/python-nlp-libraries/
[18] mail-cleaner - PyPI, https://pypi.org/project/mail-cleaner/
[19] Data Preprocessing in Machine Learning: Steps & Best Practices - lakeFS, https://lakefs.io/blog/data-preprocessing-in-machine-learning/
[20] The Effect of Text Preprocessing Strategies on Detecting Fake Consumer Reviews - Sci-Hub, https://sci-hub.se/downloads/2021-06-15/b1/barushka2019.pdf
[21] How NLP-Based Systems Combat Fraud | Insurance Thought Leadership, https://www.insurancethoughtleadership.com/claims/how-nlp-based-systems-combat-fraud
[22] Comparative Analysis of Machine Learning Models for Real-Time Fraud Detection, https://www.researchgate.net/publication/388273168_Comparative_Analysis_of_Machine_Learning_Models_for_Real-Time_Fraud_Detection
[23] Evaluating the Performance of BERT, XGBoost, and Hybrid Models for Fake News Detection, https://www.researchgate.net/publication/389499023_Evaluating_the_Performance_of_BERT_XGBoost_and_Hybrid_Models_for_Fake_News_Detection
[24] Evaluating Deep Learning vs Traditional Machine Learning Models for Real-Time Fraud Detection in Financial Systems - ResearchGate, https://www.researchgate.net/publication/390955775_Evaluating_Deep_Learning_vs_Traditional_Machine_Learning_Models_for_Real-Time_Fraud_Detection_in_Financial_Systems
[25] Application of Deep Learning in Financial Credit Card Fraud Detection - ResearchGate, https://www.researchgate.net/publication/385863027_Application_of_Deep_Learning_in_Financial_Credit_Card_Fraud_Detection
[26] Transformer-Based Fraud Detection - Abinash Kumar Mishra, https://hustlercoder.substack.com/p/transformer-based-fraud-detection-f18
[27] When to use deep learning for bot and fraud detection and when not to use it?, https://ai.stackexchange.com/questions/47279/when-to-use-deep-learning-for-bot-and-fraud-detection-and-when-not-to-use-it
[28] The Future Of Fraud Detection: Trends And Challenges - Financial Crime Academy, https://financialcrimeacademy.org/the-future-of-fraud-detection/
[29] Artificial intelligence in fraud prevention: Exploring techniques and applications challenges and opportunities - ResearchGate, https://www.researchgate.net/publication/383264952_Artificial_intelligence_in_fraud_prevention_Exploring_techniques_and_applications_challenges_and_opportunities
[30] How does multimodal AI improve fraud detection? - Milvus, https://milvus.io/ai-quick-reference/how-does-multimodal-ai-improve-fraud-detection
[31] How does multimodal AI improve fraud detection? - Zilliz Vector Database, https://zilliz.com/ai-faq/how-does-multimodal-ai-improve-fraud-detection
[32] Building a Multimodal Classifier in PyTorch: A Step-by-Step Guide | by Arpan Roy - Medium, https://medium.com/@arpanroy_43094/building-a-multimodal-classifier-in-pytorch-a-step-by-step-guide-a6dbd9900802
[33] TeleAntiFraud-28k: An Audio-Text Slow-Thinking Dataset for Telecom Fraud Detection - arXiv, https://arxiv.org/html/2503.24115v4
[34] TeleAntiFraud-28k: An Audio-Text Slow-Thinking Dataset for Telecom Fraud Detection, https://arxiv.org/html/2503.24115v2

[35] Mel-frequency cepstrum - Wikipedia, https://en.wikipedia.org/wiki/Mel-frequency_cepstrum
[36] Mel Frequency Cepstral Coefficient (MFCC) tutorial - Practical Cryptography, http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/
[37] Pitch Imperfect: Detecting Audio Deepfakes Through Acoustic Prosodic Analysis - arXiv, https://arxiv.org/abs/2502.14726
[38] Pitch Imperfect: Detecting Audio Deepfakes Through Acoustic Prosodic Analysis, https://deepfake-total.com/related_work/2502.14726
[39] Improving Short Utterance Anti-Spoofing with AASIST2 - arXiv, https://arxiv.org/html/2309.08279v2
[40] Explore wav2vec 2.0 for Mispronunciation Detection - ISCA Archive, https://www.isca-archive.org/interspeech_2021/xu21k_interspeech.pdf
[41] Siamese Network with wav2vec Feature for Spoofing Speech Detection - ResearchGate, https://www.researchgate.net/publication/354221465_Siamese_Network_with_wav2vec_Feature_for_Spoofing_Speech_Detection
[42] Anti-Spoofing Using Transfer Learning with Variational Information Bottleneck, https://www.researchgate.net/publication/363646085_Anti-Spoofing_Using_Transfer_Learning_with_Variational_Information_Bottleneck
[43] (PDF) Developing a Multimodal AI Framework for Real-Time ..., https://www.researchgate.net/publication/392263081_Developing_a_Multimodal_AI_Framework_for_Real-Time_Document_Verification_and_Fraud_Detection_Using_Cross-Modal_Feature_Fusion
[44] A Comparative Analysis of Three Data Fusion Methods and Construction of the Fusion Method Selection Paradigm - MDPI, https://www.mdpi.com/2227-7390/13/8/1218
[45] (HELP) Multimodal (Image + Audio) neural networks : r/deeplearning - Reddit, https://www.reddit.com/r/deeplearning/comments/1ig1wfg/help_multimodal_image_audio_neural_networks/
[46] Safeguarding Brand and Platform Credibility Through AI-Based Multi ..., https://www.mdpi.com/1999-5903/17/9/391
[47] Cross attention for Text and Image Multimodal data fusion - Stanford University, https://web.stanford.edu/class/cs224n/final-reports/256711050.pdf
[48] Multimodal Recommendation System Based on Cross Self-Attention Fusion - MDPI, https://www.mdpi.com/2079-8954/13/1/57
[49] [Literature Review] TeleAntiFraud-28k: A Audio-Text Slow-Thinking Dataset for Telecom Fraud Detection - Moonlight, https://www.themoonlight.io/en/review/teleantifraud-28k-a-audio-text-slow-thinking-dataset-for-telecom-fraud-detection
[50] [Literature Review] TeleAntiFraud-28k: An Audio-Text Slow-Thinking Dataset for Telecom Fraud Detection - Moonlight, https://www.themoonlight.io/en/review/teleantifraud-28k-an-audio-text-slow-thinking-dataset-for-telecom-fraud-detection
[51] What is data drift in ML, and how to detect and handle it - Evidently AI, https://www.evidentlyai.com/ml-in-production/data-drift
[52] Model Drift: Types, Causes and Early Detection - Lumenova AI, https://www.lumenova.ai/blog/model-drift-concept-drift-introduction/
[53] Real-Time Fraud Detection at Scale: Leveraging MLOps for Financial Services - TechAhead, https://www.techaheadcorp.com/blog/mlops-for-real-time-fraud-detection-for-financial-services/
[54] Experiment Tracking in Machine Learning – Everything You Need to Know - Viso Suite, https://viso.ai/deep-learning/experiment-tracking/
[55] 13 Best Tools for ML Experiment Tracking and Management in 2025, https://neptune.ai/blog/best-ml-experiment-tracking-tools
[56] Data drift case studies: Navigating the challenges in machine learning - BytePlus, https://www.byteplus.com/en/topic/400230
[57] Data drift detection and mitigation: A comprehensive MLOps approach for real-time systems, https://www.researchgate.net/publication/388187259_Data_drift_detection_and_mitigation_A_comprehensive_MLOps_approach_for_real-time_systems
[58] How to Manage AI Model Drift in FinTech Applications, https://www.fintechweekly.com/magazine/articles/ai-model-drift-management-fintech-applications

[59] Risk Mitigation Strategies - LexisNexis Risk Solutions, https://risk.lexisnexis.com/global/en/insights-resources/article/single-vs-multi-models