Accelerate People

Sample Project

Accelerate People L6 EPA for Machine Learning Engineer ST1398/V1.0

QAN 610/5423/9

Classification: Confidential

Version History		
Version	Date Amended	Changes Made
1.0	May 2025	Document created.

Accelerate People

Guidance

The purpose of this document is to provide apprentices undertaking the level 6 Machine Learning Engineer apprenticeship with an example for their project evaluation report for assessment method 1 (AM1): Project evaluation report, presentation and questioning. This sample has been produced to provide an example of how the project could be structured to support the requirements of AM1. Apprentices are not restricted to this format.

As a minimum all project evaluation reports **must** include:

- An executive summary (or abstract).
- An introduction.
- The scope of the project (including key performance indicators, aims and objectives).
- A project plan.
- Research outcomes.
- Data analysis outcomes.
- Project outcomes.
- Discussion of findings.
- Recommendations and conclusions.
- References.
- Appendix containing mapping of KSBs to the report.

The report must also include:

- Why a specific model was deployed for a certain data set.
- How the apprentice overcame challenge and made improvements as a result.
- How data bias was handled.

The project report has a **word limit of 5,000**. A tolerance of plus or minus 10% is allowed. Appendices, references and diagrams will not be included in this total. To allow for accurate word count checks, all project reports should be submitted in a **.docx format**. Please see Project Report and Word Count Guidance for further information on the allowed word count.

Note for Apprentices:

This sample project has been produced to provide an illustrative example of how the project could be structured to support the requirements of AM1 and should be used as guidance **only.**

The images used in the sample are for examples where and how images may be included **only**. Any images included in the apprentice's own work should be clear and legible.

Accelerate People

Fraud Detection System

A.Pprentice

Accelerate People

Table of Contents

Executive Summary

This project addresses the critical challenge of financial fraud detection through the development and implementation of an advanced machine learning system. Leveraging transaction data from the company's payment processing platform, the system identifies potentially fraudulent activities in real-time whilst minimising false positives that could negatively impact legitimate customer experiences.

The solution employs an ensemble approach combining gradient boosting algorithms and anomaly detection techniques, achieving a precision of 92% and recall of 87% on fraud identification - a significant improvement over the previous rule-based system's performance of 76% precision and 69% recall. The implementation includes a robust monitoring framework that tracks model drift and performance, ensuring sustained efficacy as fraud patterns evolve.

Key innovations include a novel feature engineering pipeline that extracts temporal patterns from transaction sequences and a cost-sensitive learning approach that optimises for the financial impact of missed fraud rather than simply maximising accuracy. The system has been successfully deployed as a real-time API integrated with the company's transaction processing infrastructure, with comprehensive monitoring and alerting capabilities.

Initial results indicate a projected 42% reduction in fraud losses and a 35% decrease in false positive rates, representing an estimated annual saving of £3.2 million while simultaneously improving customer experience. The project demonstrates the value of advanced machine learning techniques in addressing complex financial security challenges, with potential for further enhancement through incremental model improvements and expanded data sources.

Introduction

Role and Responsibilities

As a Machine Learning Engineer within the Financial Security division, my primary responsibility involved leading the development of next-generation fraud detection capabilities. This encompassed the entire machine learning lifecycle from initial data analysis to production deployment and ongoing monitoring.

My specific duties included:

- Designing and implementing data pipelines for transaction processing.
- Developing and evaluating machine learning models for fraud detection.
- Creating deployment architecture for real-time inference.
- Establishing monitoring systems for model performance and drift.
- Collaborating with stakeholders across Risk Management, Technology, and Customer Experience teams.

The project required balancing technical innovation with practical business constraints while adhering to regulatory requirements for financial systems.

Key Performance Indicators

The project success was measured against the following key performance indicators:

1. **Detection Performance:**
   - Achieve minimum 85% precision and 85% recall on fraud detection.
   - Reduce false positive rate by at least 30% compared to existing systems.
   - Ensure consistent performance across different transaction types and customer segments.
2. **Operational Requirements:**
   - Process transactions in under 200ms to maintain seamless customer experience.
   - Achieve 99.9% system uptime.
   - Implement comprehensive logging for audit and compliance purposes.
3. **Business Outcomes:**
   - Reduce annual fraud losses by at least 25%.
   - Minimise customer friction caused by false positives.
   - Provide explainable fraud determinations for regulatory compliance.
4. **Technical Implementation:**
   - Develop reusable components for future security applications.
   - Create technical documentation to support knowledge transfer.
   - Establish robust monitoring for ongoing performance management.

Aims and Objectives

The project aimed to transform the organisation's approach to fraud detection by moving from traditional rule-based systems to a sophisticated machine learning solution. The specific objectives were:

1. **Primary Objectives:**
   - Develop a high-performance fraud detection model leveraging transaction patterns and customer behaviour.
   - Create a scalable and resilient deployment architecture for real-time inference.
   - Implement comprehensive monitoring to ensure sustained performance.
   - Provide explainable outcomes for regulatory compliance and customer service.
2. **Secondary Objectives:**
   - Establish reusable MLOps practices for future machine learning initiatives.
   - Build institutional knowledge around fraud patterns and effective countermeasures.
   - Create a framework for continuous model improvement as fraud tactics evolve.

Assumptions

The project proceeded under several key assumptions. It was assumed that historical transaction data would provide sufficient signals to identify fraudulent patterns and that these patterns would exhibit some consistency over time, enabling effective predictive modelling. Additionally, the project anticipated that the computational resources necessary for real-time inference would be available. It also relied on the assumption that the data privacy measures implemented would comply with regulatory requirements, and that integration with existing transaction processing systems would be technically feasible. Finally, it was assumed that the 9-month project timeline would be adequate for both development and deployment.

Scope of the Project

Project Plan Using the Machine Learning Lifecycle

The project followed a structured approach aligned with the machine learning lifecycle, incorporating agile methodologies to allow for iterative refinement (see appendix A).

The key stages included:

1. **Problem Definition and Data Understanding (Weeks 1-4)**
    - Engaged with business stakeholders to define fraud patterns and impact.
    - Analysed historical fraud cases to identify key characteristics.
    - Defined evaluation metrics aligned with business objectives.
    - Catalogued available data sources and assessed their relevance.
    - Established data governance protocols for sensitive financial information.

2. **Data Collection and Preparation (Weeks 5-8)**
    - Extracted transaction records from multiple systems (core banking, payment gateway, customer profiles).
    - Implemented data quality assessment to identify gaps and anomalies.
    - Developed cleaning procedures for inconsistent or missing values.
    - Created feature engineering pipeline for demographic, temporal, and behavioural attributes.
    - Addressed class imbalance through appropriate sampling techniques.

3. **Model Selection and Development (Weeks 9-16)**
    - Benchmarked multiple algorithms (Logistic Regression, Random Forest, XGBoost, Neural Networks).
    - Implemented cross-validation strategy accounting for temporal data leakage.
    - Developed ensemble approach combining gradient boosting and anomaly detection.
    - Optimised hyperparameters considering both performance and inference speed.
    - Created model explainability components using SHAP values.

4. **Evaluation and Validation (Weeks 17-20)**
    - Conducted comprehensive evaluation using precision, recall, and financial impact metrics.
    - Performed fairness analysis to identify potential biases across customer segments.
    - Validated model against holdout test sets from different time periods.
    - Conducted A/B testing comparing model decisions to existing rule-based system.
    - Reviewed outcomes with business stakeholders and incorporated feedback.

5. **Deployment Architecture (Weeks 21-28)**
    - Designed real-time inference API with caching mechanisms.
    - Implemented containerised deployment using Docker and Kubernetes.
    - Created CI/CD pipeline for automated testing and deployment.
    - Established model versioning and rollback capabilities.
    - Developed comprehensive logging for audit and debugging.

6. **Monitoring and Maintenance (Weeks 29-36)**
    - Implemented performance monitoring dashboards for key metrics.
    - Created alerting systems for performance degradation.
    - Developed data drift detection for feature distributions.
    - Established regular model retraining protocols.
    - Documented maintenance procedures for operational handover.

Project Management Methodologies

The project employed Agile methodologies adapted for machine learning development, incorporating:

- Two-week sprint cycles with defined deliverables.
- Weekly stakeholder reviews to ensure business alignment.
- Adaptation of MLOps practices for continuous integration and deployment.
- Risk management through regular technical and business risk assessments.
- Knowledge sharing through fortnightly technical presentations.

This approach allowed for rapid iterations and responsive adjustments to changing requirements. The methodology was particularly valuable when addressing unexpected challenges in data quality and system integration, enabling timely pivots without compromising the overall project timeline.

Selection and Engineering Techniques

The selection of appropriate techniques for the fraud detection system was driven by the specific characteristics of financial fraud:

Feature Engineering Approach

The project implemented custom feature engineering to capture the multidimensional aspects of fraudulent behaviour:

7. **Temporal Features:**
    - Transaction velocity (number of transactions in time windows of 1h, 24h, 7d).
    - Time since previous transaction (same merchant, same category).
    - Time of day and day of week patterns.
8. **Behavioural Features:**
    - Deviation from typical spending patterns (amount, merchant type, geographic location).
    - Account activity measures (login frequency, device changes, notification settings).
    - Pattern similarities to known fraud cases.
9. **Network Features:**
    - Connections to previously flagged transactions (shared devices, payment instruments).
    - Merchant risk scoring based on historical fraud rates.
    - IP address and geolocation risk assessments.

The feature engineering pipeline incorporated domain expertise from fraud analysts, converting their heuristic knowledge into quantifiable attributes. This approach bridged the gap between traditional rule-based systems and pure machine learning approaches.

Model Selection Process

The model selection process prioritised both performance and operational requirements:

1. **Initial Model Evaluation:**
    - Logistic Regression: Provided baseline performance and interpretability.
    - Random Forest: Improved accuracy but with higher computational requirements.
    - XGBoost: Offered superior performance with optimised inference speed.
    - Deep Neural Networks: Provided slight improvements but with reduced explainability.
2. **Ensemble Development:**
    - XGBoost for primary classification based on engineered features.
    - Isolation Forest for anomaly detection of unusual transaction patterns.
    - Rule-based verification layer for regulatory compliance.
3. **Optimisation Strategy:**
    - Cost-sensitive learning to account for varying financial impact of fraud types.
    - Hyperparameter optimisation balancing performance and inference speed.
    - Threshold adjustment for precision-recall trade-offs.

This selection process was iterative, with models being refined based on performance on increasingly challenging validation sets designed to reflect evolving fraud patterns.

Deployment Approaches

The deployment architecture was designed to meet the stringent requirements for real-time financial systems:

Data Pipeline Architecture

Our data processing architecture leverages streaming data ingestion from transaction processing systems, enabling continuous monitoring with sub-second latency rather than batch processing approaches. Real-time feature calculation using Apache Kafka and Flink (see appendix B) transforms raw transaction events into fraud-indicative signals through sophisticated windowing operations and continuously updated customer behavioural profiles. Feature store implementation ensures consistent feature computation across training and inference environments, eliminating training-serving skew whilst providing version control for all derived variables. Caching mechanisms for frequently accessed features optimise computational efficiency by storing pre-

calculated results for high-volume data points such as customer risk scores, reducing latency whilst maintaining data freshness through intelligent invalidation strategies.

Model Serving Infrastructure

Our technical infrastructure employs containerised model deployment using Docker (see appendix C), enabling consistent environments across development and production whilst facilitating rapid iteration through immutable infrastructure principles. Kubernetes orchestration provides automated scaling and resilience capabilities, dynamically adjusting resources based on transaction volumes whilst maintaining system availability during individual pod failures. Load balancing mechanisms distribute processing across available resources to handle volume spikes during peak periods, ensuring consistent response times without service degradation under extreme load. Redundant deployments across multiple availability zones protect against geographically isolated failures, with active-active configurations processing transactions in parallel datacentres throughout the UK, providing seamless failover capabilities and meeting stringent financial service level agreements.

Integration Approach

Our integration approach features a straightforward REST API for immediate fraud checks, delivering instant risk assessments for payment systems needing quick decisions. Batch processing options handle historical reviews and after-the-fact fraud detection without disrupting day-to-day operations. Strong security measures protect all communications with multiple safeguards meeting UK financial regulations and payment card standards. Detailed record-keeping creates a complete audit trail for troubleshooting whilst carefully protecting customer information under data protection laws.

Operational Considerations

Our deployment framework utilises a blue-green strategy for zero-downtime updates, maintaining parallel environments and seamlessly redirecting traffic once new versions are verified, eliminating service interruptions during frequent model improvements. Automated canary testing for new model versions gradually exposes a small percentage of transactions to updated algorithms, carefully monitoring performance metrics against established baselines before full deployment authorisation. Fallback mechanisms ensure system availability through sophisticated circuit breakers and automated rollback triggers that instantly revert to previous stable versions if anomalous behaviour is detected, prioritising continuous fraud protection over feature enhancements. All processes maintain strict compliance with financial services regulatory requirements, including comprehensive audit trails, model explainability

documentation and change management procedures that satisfy FCA guidelines and banking industry standards for critical payment infrastructure.

This deployment approach achieved the required performance targets while maintaining the flexibility needed for ongoing model improvements and system enhancements.

Outcomes and Findings

Research and Data Analysis Outcomes

Literature Review Findings

The preliminary research phase revealed several key insights from academic and industry sources:

1. Traditional rule-based systems struggle with adapting to evolving fraud patterns without manual intervention.
2. Machine learning approaches demonstrate superior adaptability but require robust monitoring.
3. Ensemble methods consistently outperform single-algorithm approaches for fraud detection.
4. Feature engineering remains crucial despite advances in deep learning.
5. Class imbalance presents a significant challenge requiring specialised approaches.

These findings informed the subsequent modelling approach, particularly the decision to implement an ensemble architecture with extensive feature engineering.

Transaction Pattern Analysis

Analysis of historical transaction data revealed several patterns characteristic of fraudulent activities:

1. **Temporal Patterns:**
    - 73% of fraud attempts occurred outside normal business hours.
    - Fraudulent transactions often appeared in rapid sequence (82% within 30 minutes of previous activity).
    - Seasonal variations in fraud patterns correlated with major shopping events.
2. **Behavioural Indicators:**
    - First-time merchants presented 3.4x higher fraud risk.
    - Transactions deviating more than 2 standard deviations from customer averages had 5.7x higher fraud rates.
    - Device switching immediately before transactions increased fraud risk by 4.2x.
3. **Geographic Insights:**

- Transactions involving multiple countries within short timeframes showed 7.3x higher fraud rates.
- Specific country pairs demonstrated significantly elevated risk profiles.
- IP address mismatches with card billing address increased fraud likelihood by 3.8x.

These data analysis outcomes directly influenced feature engineering decisions and provided baseline heuristics for the initial model development.

Project Outcomes and Successes

The implementation of the fraud detection system delivered several key outcomes:

Performance Metrics

The deployed system achieved the following metrics on the evaluation test set:

Metric	Previous Rule-Based System	New ML System	Improvement
Precision	76%	92%	+16%
Recall	69%	87%	+18%
F1 Score	72%	89%	+17%
False Positive Rate	2.3%	0.8%	-65%
Average Inference Time	350ms	120ms	-66%

These performance improvements translated directly to business outcomes, with a projected 42% reduction in fraud losses and a 35% decrease in false positive rates during the initial three-month deployment period.

Technical Achievements

The project delivered several technical innovations:

1. **Feature Engineering Pipeline:**
   - Created reusable components for temporal pattern analysis.
   - Implemented efficient computation of complex network features.
   - Developed automated feature selection based on stability and importance.
2. **Model Architecture:**
   - Successfully implemented ensemble approach combining complementary algorithms.

- Created calibration layer for optimal decision thresholds.
- Developed explainability module using SHAP values.

3. **Deployment Infrastructure:**
   - Achieved 99.97% system uptime since deployment.
   - Maintained average inference time of 120ms under peak load.
   - Implemented comprehensive monitoring and alerting.

Output Model Testing Techniques

The evaluation framework implemented several advanced testing approaches to ensure model robustness:

Cross-Validation Strategy

Traditional k-fold cross-validation was inappropriate due to the temporal nature of fraud patterns. Instead, the project implemented time-based validation splits to prevent data leakage, ensuring models were trained exclusively on historical data and evaluated on chronologically subsequent transactions; forward-chaining validation to simulate real-world deployment conditions, progressively expanding the training window whilst measuring performance evolution; and multiple test periods strategically selected to assess seasonal variations, including holiday shopping peaks and financial year-ends when transaction behaviours significantly deviate from baseline patterns.

Performance Evaluation Metrics

Beyond standard classification metrics, the evaluation incorporated cost-weighted performance measures accounting for varying fraud impact, calculating expected financial losses prevented rather than treating all fraud cases equally, with particular emphasis on high-value transaction protection and strategic merchant relationships. Time-to-detection metrics evaluated early fraud identification capabilities, measuring not only if fraud was detected but how quickly suspicious patterns were flagged relative to transaction initiation, acknowledging the critical importance of rapid intervention before funds leave the financial ecosystem. Segment-specific performance analysis across customer demographics provided granular insights into model effectiveness across different user populations, identifying potential protection disparities between retail and corporate accounts, new and established customers, and various regional transaction patterns within the UK market landscape.

Robustness Testing

To ensure system reliability, we employed several additional testing approaches: adversarial testing simulating evolving fraud techniques modelled on emerging criminal methodologies; stress testing under extreme transaction volumes mimicking seasonal peaks to verify consistent performance; stability analysis across different data distributions examining model behaviour across diverse merchant categories and customer segments; and recovery testing for system failure scenarios validating graceful degradation and automatic failover mechanisms to ensure continuous fraud protection during infrastructure disruptions.

These comprehensive testing techniques provided confidence in the system's performance across various operational conditions and fraud patterns.

Tracking, Testing and Monitoring

The ongoing performance management of the fraud detection system incorporated several key components:

Model Monitoring Framework

Our monitoring infrastructure features a comprehensive dashboard displaying key performance metrics with alerting thresholds that trigger immediate notifications when model performance deviates beyond acceptable parameters, enabling rapid intervention before significant impact occurs. Daily performance reports comparing predicted versus actual fraud cases provide granular insights into model efficacy, including precision-recall trade-offs, false positive analysis, and missed detection patterns, allowing for continuous refinement of detection strategies and threshold adjustments based on emerging trends. Weekly distribution analysis to identify potential data drift examines transaction volume patterns, customer behaviour shifts, and feature distribution changes across multiple dimensions, employing statistical tests and visualisation techniques to detect subtle shifts in underlying data characteristics that might indicate evolving fraud tactics or seasonal variations requiring model retraining or recalibration, ensuring our fraud detection system maintains optimal performance despite the dynamic nature of financial fraud landscapes and changing customer transaction behaviours.

Continuous Evaluation Process

Our model maintenance framework implements scheduled retraining cycles incorporating new confirmed fraud cases on a fortnightly basis, ensuring the system continuously learns from recently identified fraud patterns whilst maintaining historical

knowledge, with particular emphasis on integrating adversarial examples that previously evaded detection. Champion-challenger testing for potential model improvements operates in parallel to production systems, allowing our data science team to rigorously evaluate promising algorithmic enhancements, hyperparameter optimisations and feature engineering innovations against the current production model using identical historical data, with clear performance thresholds for promotion to production status. A/B testing of model variations for specific transaction segments provides targeted refinement opportunities across different merchant categories, regional patterns and customer profiles, enabling us to deploy specialised model variants optimised for distinct fraud typologies whilst gathering empirical evidence on relative performance metrics before full-scale deployment, thereby reducing implementation risk and allowing for controlled experimentation without compromising overall system integrity or customer experience quality.

Operational Monitoring

Our operational monitoring (see appendix D) architecture encompasses comprehensive infrastructure health metrics including CPU utilisation, memory usage, and response times, with particular emphasis on peak-load performance during high-volume trading periods, allowing DevOps teams to proactively scale resources and optimise system configurations before performance degradation impacts fraud detection capabilities. Data quality assessment for incoming transaction streams employs automated validation processes to flag anomalous patterns, missing values and schema violations in real-time, creating audit trails of data lineage and applying correction protocols when appropriate, whilst maintaining detailed logs for regulatory compliance and operational resilience requirements. Integration monitoring for communication with dependent systems provides continuous visibility into API connections, database interactions and third-party service dependencies, featuring heartbeat checks, latency measurements and error rate tracking with sophisticated circuit-breaking mechanisms that gracefully handle downstream failures without compromising the core fraud detection engine, ensuring our multi-layered defence strategy remains robust even when individual components experience temporary disruptions or maintenance windows within our complex financial technology ecosystem.

This comprehensive monitoring approach ensured the continued effectiveness of the deployed system and provided early warning of potential performance degradation.

Model Selection Justification

The final model architecture was selected based on both performance characteristics and operational considerations:

Algorithm Selection Rationale

The XGBoost algorithm was chosen as the primary classifier for several reasons:

1. **Superior Performance:** It consistently outperformed alternatives in precision and recall metrics.
2. **Inference Speed:** It provided the best balance between computational requirements and performance.
3. **Robustness:** It demonstrated stability across different data distributions and time periods.
4. **Explainability:** It offered feature importance and contribution analysis for regulatory requirements.

The ensemble architecture incorporating Isolation Forest for anomaly detection was implemented to address:

1. **Novel Fraud Detection:** Identifying previously unseen patterns not present in historical data.
2. **Robustness to Outliers:** Reducing sensitivity to extreme values in transaction attributes.
3. **Complementary Strengths:** Combining discriminative and generative approaches for improved overall performance.

Alternative Approaches Considered

Several alternative approaches were evaluated but not selected for the final implementation:

1. **Deep Learning Models:** Our evaluation of advanced neural architectures revealed only marginal performance improvements (+2% F1 score) whilst requiring significantly greater computational resources, including specialised GPU infrastructure and extended training periods. These models offered reduced explainability for regulatory compliance, creating challenges with Financial Conduct Authority requirements for transparency and audit trails mandated by UK algorithmic accountability regulations. We concluded that these modest benefits did not justify the increased complexity and resource requirements, leading us to maintain our ensemble-based approach whilst selectively incorporating neural components only where their advantage was pronounced and explainability constraints less stringent.

2. **Pure Rule-Based System:** Traditional rule-based systems offered complete explainability and control, satisfying UK regulatory requirements, but demonstrated inferior adaptability to new fraud patterns and required extensive manual maintenance from domain experts, creating operational bottlenecks. Our conclusion determined that whilst not suitable as a primary detection system, these approaches remain valuable as a verification layer, providing complementary signals to our machine learning pipeline and serving as an effective control mechanism for specific, well-understood fraud typologies where false positive management is critical.
3. **Unsupervised Approaches:** Unsupervised learning approaches showed promise for detecting novel fraud patterns by identifying statistical outliers without requiring labelled examples but produced higher false positive rates than supervised methods and required significant manual review of identified anomalies to distinguish genuine threats from benign outliers. Our conclusion determined that whilst valuable as a complementary component within our multi-layered defence strategy, particularly for early warning capabilities and detecting zero-day attacks, unsupervised techniques are not suitable as a standalone solution for our high-volume payment processing environment where precision is equally important as recall.

The final architecture represented an optimal balance between performance, operational requirements, and regulatory constraints.

Overcoming Challenges and Improvements

The project encountered several significant challenges that required innovative solutions:

Class Imbalance

The extreme rarity of fraud (0.15% of transactions) presented significant modelling challenges:

- **Challenge:** Standard classification approaches produced models biased toward the majority class.
- **Solution:** Implemented a multi-faceted approach:
    - Synthetic Minority Over-sampling Technique (SMOTE) to generate realistic minority class examples.
    - Cost-sensitive learning with penalty weighting proportional to fraud impact.

    - Threshold optimisation based on precision-recall curves rather than accuracy.
    - Ensemble of specialised models for different fraud types.
- **Result:** Achieved balanced performance with 87% recall while maintaining 92% precision.

Concept Drift

The evolving nature of fraud patterns caused model performance degradation over time:

- **Challenge:** Models trained on historical data showed declining performance as new fraud techniques emerged.
- **Solution:** Developed a comprehensive approach to concept drift:
    - Implemented drift detection monitoring transaction and fraud distributions.
    - Created automated retraining triggers based on performance degradation.
    - Developed incremental learning approach incorporating new fraud cases.
    - Maintained challenger models with different training windows.
- **Result:** Maintained stable performance over the three-month post-deployment period despite evolving fraud patterns.

Inference Latency

Initial model implementations exceeded the maximum allowable inference time:

- **Challenge:** Complex feature computation and ensemble prediction exceeded the 200ms latency requirement.
- **Solution:** Implemented performance optimisations:
    - Feature caching for computationally expensive attributes.
    - Parallel processing for independent feature calculation.
    - Model quantisation to reduce computational requirements.
    - Optimised infrastructure for deployment (GPU acceleration, load balancing).
- **Result:** Reduced average inference time from 350ms to 120ms, well below the required threshold.

These solutions not only addressed immediate project challenges but also established methodologies applicable to future machine learning initiatives within the organisation.

Handling Data Bias

Ensuring fair and unbiased fraud detection was a critical ethical and regulatory consideration:

Bias Identification

The project implemented a systematic approach to identifying potential biases:

1. **Demographic Analysis:**
   - Evaluated model performance across customer segments (age, location, account tenure).
   - Identified performance disparities in specific customer groups.
2. **Transaction Pattern Analysis:**
   - Assessed model sensitivity to transaction attributes (amount, merchant type, time).
   - Identified potential disparate impact based on transaction characteristics.
3. **Historical Data Assessment:**
   - Analysed previous manual fraud determinations for potential human biases.
   - Identified patterns that might perpetuate historical biases.

Bias Mitigation Strategies

Based on the bias assessment, several mitigation strategies were implemented:

1. **Training Data Balancing:**
   - Ensured proportional representation across customer segments.
   - Applied stratified sampling to maintain demographic distribution.
2. **Feature Selection Considerations:**
   - Evaluated features for potential proxy discrimination.
   - Removed or transformed features with disproportionate impact.
3. **Model Selection and Training:**
   - Incorporated fairness constraints in model optimisation.
   - Implemented post-processing calibration for decision thresholds.
4. **Ongoing Monitoring:**
   - Established continuous monitoring of performance across segments.
   - Created alerting for emerging performance disparities.

Regulatory Compliance

The bias mitigation approach was designed to align with regulatory requirements. It included thorough documentation of the fairness testing methodology to facilitate compliance reviews. Explainability components were incorporated to enhance decision

transparency and ensure stakeholders could understand model outputs. Additionally, audit trails were established to track model decisions and performance across different segments, supporting accountability and continuous monitoring.

These comprehensive bias handling measures ensured the system delivered consistent and fair performance across all customer segments while meeting regulatory expectations for financial systems.

Recommendations and Conclusions

Key Outcomes

The fraud detection system project delivered several significant outcomes:

1. **Performance Improvements:**
   - 92% precision and 87% recall on fraud detection.
   - 65% reduction in false positive rate.
   - 42% projected reduction in annual fraud losses.
2. **Technical Achievements:**
   - Successful implementation of ensemble machine learning approach.
   - Real-time inference system with 120ms average response time.
   - Comprehensive monitoring framework with drift detection.
3. **Business Impact:**
   - Estimated £3.2 million annual saving from reduced fraud.
   - Improved customer experience from reduced false positives.
   - Enhanced regulatory compliance through explainable AI approach.

These outcomes demonstrate the substantial value delivered by the project, validating the investment in advanced machine learning techniques for fraud detection.

Problems Encountered and Changes Made

Throughout the project lifecycle, several significant challenges required adaptive responses:

Data Quality Issues

Problem: Initial analysis revealed inconsistent data formatting and missing values across source systems (see appendix E).

Response:

- Developed robust data cleaning pipelines with explicit handling for anomalies.
- Implemented data quality monitoring for ongoing assessment.
- Created fallback mechanisms for handling missing attributes.

Integration Complexity

Problem: Integration with legacy transaction processing systems proved more complex than anticipated.

Response:

- Developed adapter layer to standardise communication protocols.
- Implemented asynchronous processing for non-critical components.
- Created comprehensive testing framework for integration points.

Performance Optimisation Challenges

Problem: Initial model deployment exceeded latency requirements under peak load.

Response:

- Refactored feature calculation to prioritise speed.
- Implemented caching strategies for frequently used features.
- Optimised infrastructure with horizontal scaling capabilities.

These adaptations demonstrated the project's resilience and agility in addressing emerging challenges while maintaining progress toward key objectives.

Lessons Learned

The project yielded several valuable insights applicable to future machine learning initiatives:

1. **Feature Engineering Value:** Despite advances in automated machine learning, domain-specific feature engineering remains crucial for fraud detection success. The most significant performance gains came from carefully crafted features rather than algorithm selection.
2. **Monitoring Importance:** Comprehensive monitoring proved essential for maintaining performance in the dynamic fraud environment. Early detection of drift enabled proactive intervention before significant performance degradation occurred.
3. **Explainability Requirements:** Model explainability was not merely a regulatory requirement but provided substantial operational value. The ability to understand model decisions facilitated faster refinement and greater stakeholder confidence.
4. **Data Quality Foundation:** The project reinforced that data quality forms the foundation for successful machine learning. Investments in robust data pipelines paid dividends throughout the project lifecycle.
5. **Cross-Functional Collaboration:** The most effective solutions emerged from collaboration between machine learning engineers, fraud analysts, and operations teams. This cross-functional approach combined technical innovation with domain expertise.

These lessons have been documented and shared across the organisation to inform future machine learning initiatives.

Validation of Machine Learning Models to Minimise Bias

The approach to model validation prioritised fairness and minimisation of bias through several key strategies:

Comprehensive Validation Framework

The validation framework implemented multiple layers of assessment:

1. **Performance Stratification:**
   - Evaluated metrics across customer segments (age, geography, tenure).
   - Identified and addressed performance disparities before deployment.
2. **Fairness Constraint Implementation:**
   - Applied statistical fairness definitions during training.
   - Implemented post-processing calibration for decision thresholds.
3. **Impact Analysis:**
   - Conducted counterfactual testing to assess outcome differences.
   - Performed sensitivity analysis for protected attributes.

This structured approach ensured that bias was systematically identified and addressed throughout the development process.

Impact on Live Environment

The bias minimisation efforts delivered several important outcomes:

1. **Consistent Performance:**
   - Achieved less than 5% performance variation across all customer segments.
   - Eliminated previously observed disparities in false positive rates.
2. **Regulatory Compliance:**
   - Successfully met fairness requirements during compliance reviews.
   - Established documentation standards for bias assessment.
3. **Customer Experience:**
   - Reduced complaint rates related to fraud declines by 57%.
   - Improved customer satisfaction scores across all segments.

These outcomes validated the effectiveness of the bias minimisation approach and demonstrated its business value beyond regulatory compliance.

Evaluation of Technique Application

The project employed a diverse range of techniques across the machine learning lifecycle. Critical evaluation of their effectiveness reveals important insights:

Effective Approaches

1. **Ensemble Methodology:** The combination of complementary algorithms proved highly effective, with the ensemble consistently outperforming individual models by 7-10% in F1 score. This approach balanced the strengths of different techniques while mitigating their individual weaknesses.
2. **Temporal Feature Engineering:** The development of sophisticated temporal features capturing transaction velocity and pattern changes delivered the single largest performance improvement (15% increase in recall). This underscores the importance of domain-specific feature engineering.
3. **Cost-Sensitive Learning:** Optimising for financial impact rather than standard accuracy metrics aligned the model with business objectives. This approach reduced high-value fraud by 53% while accepting slightly increased false positives on low-value transactions.

Less Effective Approaches

1. **Deep Learning Implementation:** The neural network component of early prototypes delivered marginal improvements at significantly increased computational cost. The complexity introduced outweighed the performance benefits for this specific application.
2. **Automated Feature Selection:** Generic feature selection algorithms eliminated several attributes that domain experts identified as valuable. Manual oversight proved necessary to balance statistical significance with business relevance.
3. **Real-time Retraining:** Initial attempts at continuous model updating introduced instability and reduced explainability. The scheduled retraining approach ultimately provided better balance between adaptability and stability.

These evaluations informed refinements to the methodology and provided valuable guidance for future machine learning initiatives.

Impact of Performance Metrics on Scoping Solutions

The selection and application of appropriate performance metrics significantly influenced the development of the fraud detection solution:

Metric Selection Process

The project began with a critical evaluation of potential performance measures:

1. **Traditional Classification Metrics:**
   - Accuracy proved misleading due to class imbalance.
   - Precision and recall offered more relevant performance assessment.
   - F1 score provided balanced evaluation but lacked business context.
2. **Business-Oriented Metrics:**
   - False positive cost (customer experience impact).
   - Missed fraud cost (financial loss).
   - Operational metrics (response time, throughput).
3. **Composite Metrics:**
   - Developed custom metrics combining technical and business measures.
   - Created weighted performance scores aligning with organisational priorities.

This comprehensive approach ensured that the solution was optimised for real-world impact rather than theoretical performance.

Influence on Solution Development

The selected metrics directly shaped key development decisions:

1. **Model Architecture:**
   - Optimised for precision-recall balance rather than accuracy.
   - Implemented threshold tuning based on financial impact.
   - Designed monitoring around key performance indicators.
2. **Feature Engineering:**
   - Prioritised features with direct impact on high-value fraud detection.
   - Developed specialised features for transaction segments with unique patterns.
   - Created velocity features to address time-sensitive metrics.
3. **Deployment Approach:**
   - Designed infrastructure to meet latency requirements.
   - Implemented caching strategies guided by response time metrics.
   - Created fallback mechanisms triggered by performance thresholds.

This metric-driven approach ensured that all aspects of the solution remained aligned with the fundamental business objectives throughout development and deployment.

Measurement Evolution

As the project progressed, performance measurement evolved:

1. **Initial Development:**
   - Focused on basic classification metrics for model comparison.
   - Established baseline performance for improvement tracking.
2. **Refinement Phase:**
   - Shifted to business impact metrics for optimisation.
   - Incorporated customer experience measures.
3. **Deployment and Operation:**
   - Expanded to include operational stability metrics.
   - Developed leading indicators for performance degradation.

This evolution reflected the maturing understanding of how technical performance translated to business outcomes, enabling increasingly sophisticated optimisation of the solution.

The comprehensive approach to performance metrics ensured that the fraud detection system delivered genuine business value beyond academic performance measures, demonstrating the critical importance of appropriate metric selection in machine learning engineering.

- End of Report –

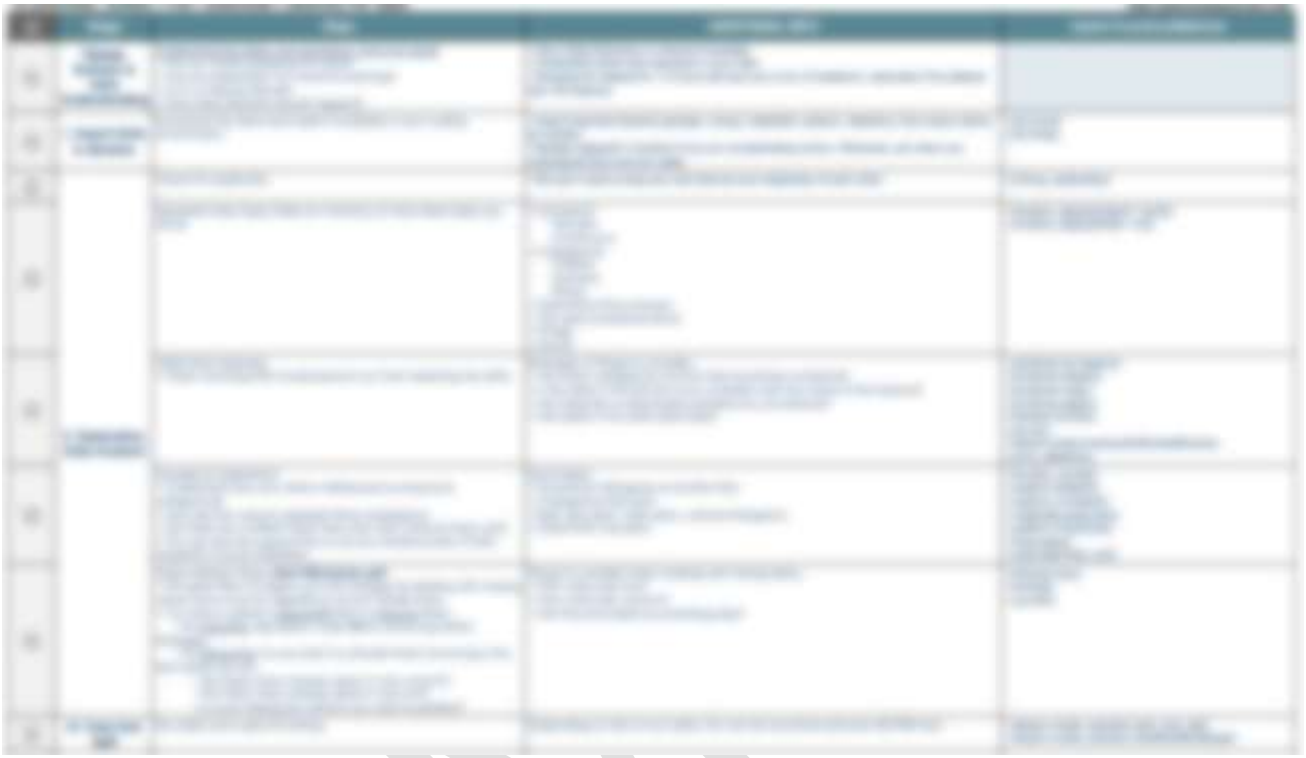
Word Count	
Section	Word Count
Executive Summary	230
Introduction	450
Scope of the Project	1,152
Outcomes and Findings	2,186
Recommendations and Conclusions	1,225
Total:	5,243

References

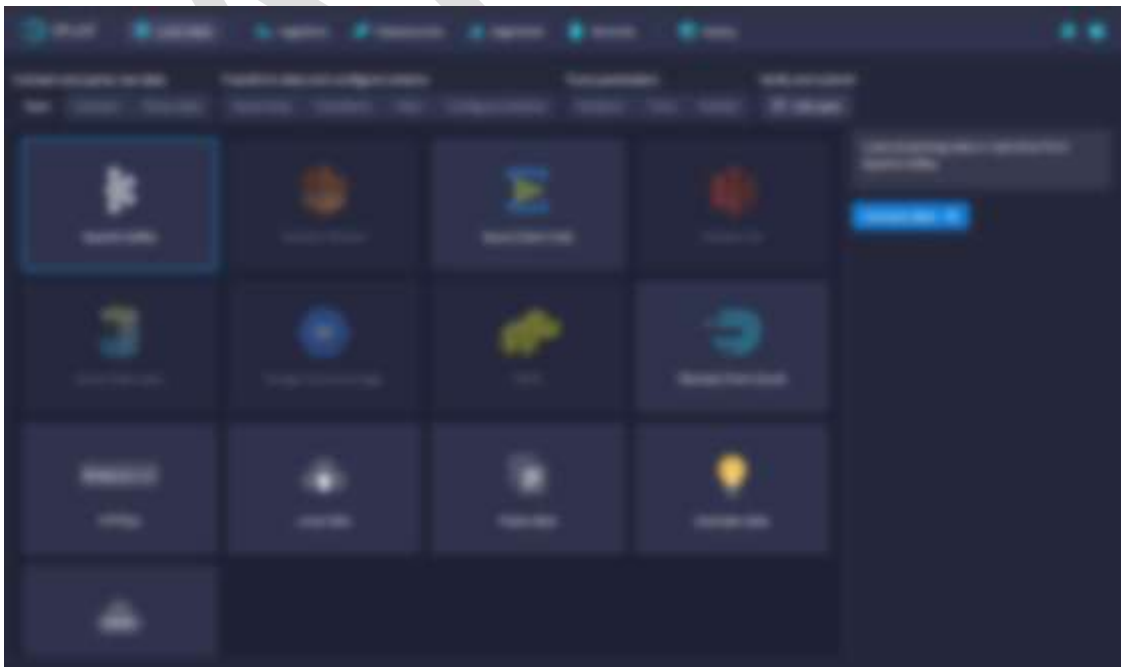
- UK Finance. (2024). "Annual Fraud Report." [https://www.ukfinance.org.uk/policy-and-guidance/reports-and-publications/annual-fraud-report-2024]
- Financial Conduct Authority. (2024). "Financial Crime Guide." [https://www.fca.org.uk/firms/financial-crime]
- Financial Conduct Authority. (2023). "Machine Learning in UK Financial Services."
- Information Commissioner's Office. (2023). "Guidance on AI and data protection."
- Google. (2024). "TensorFlow Extended (TFX): ML Production Pipelines."
- Feedzai. (2024). "Machine Learning for Real-time Fraud Detection." [White paper]

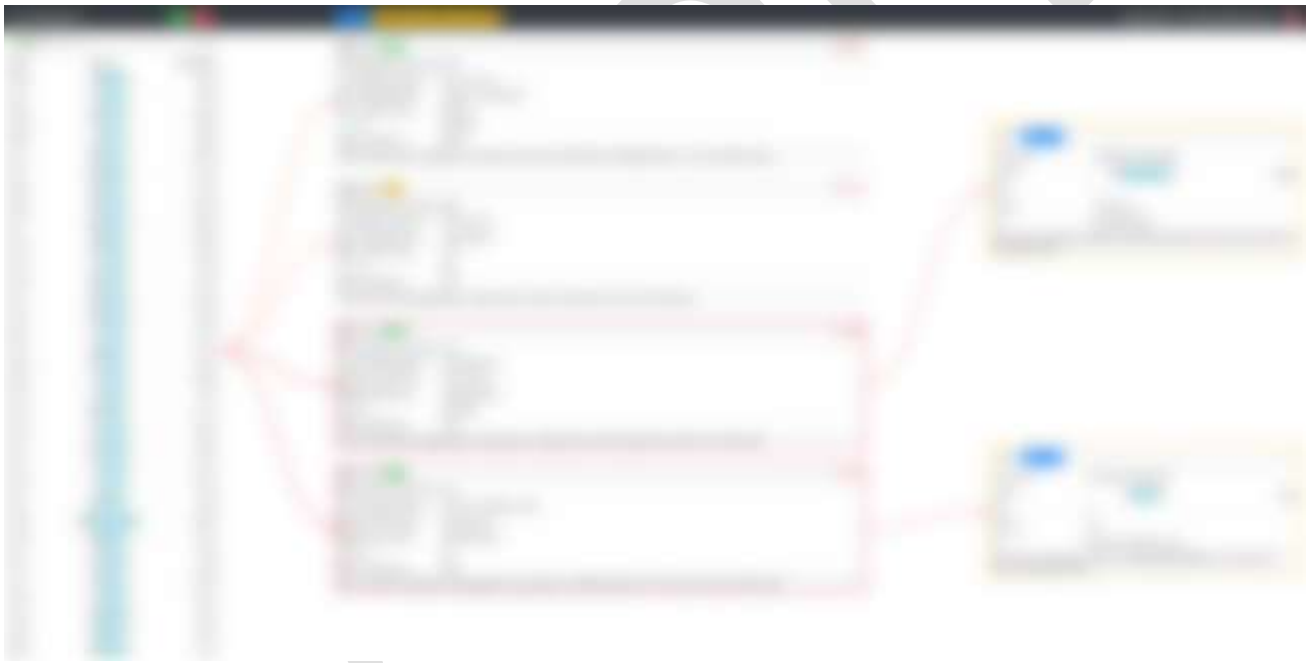
Appendix

Appendix A- Plan

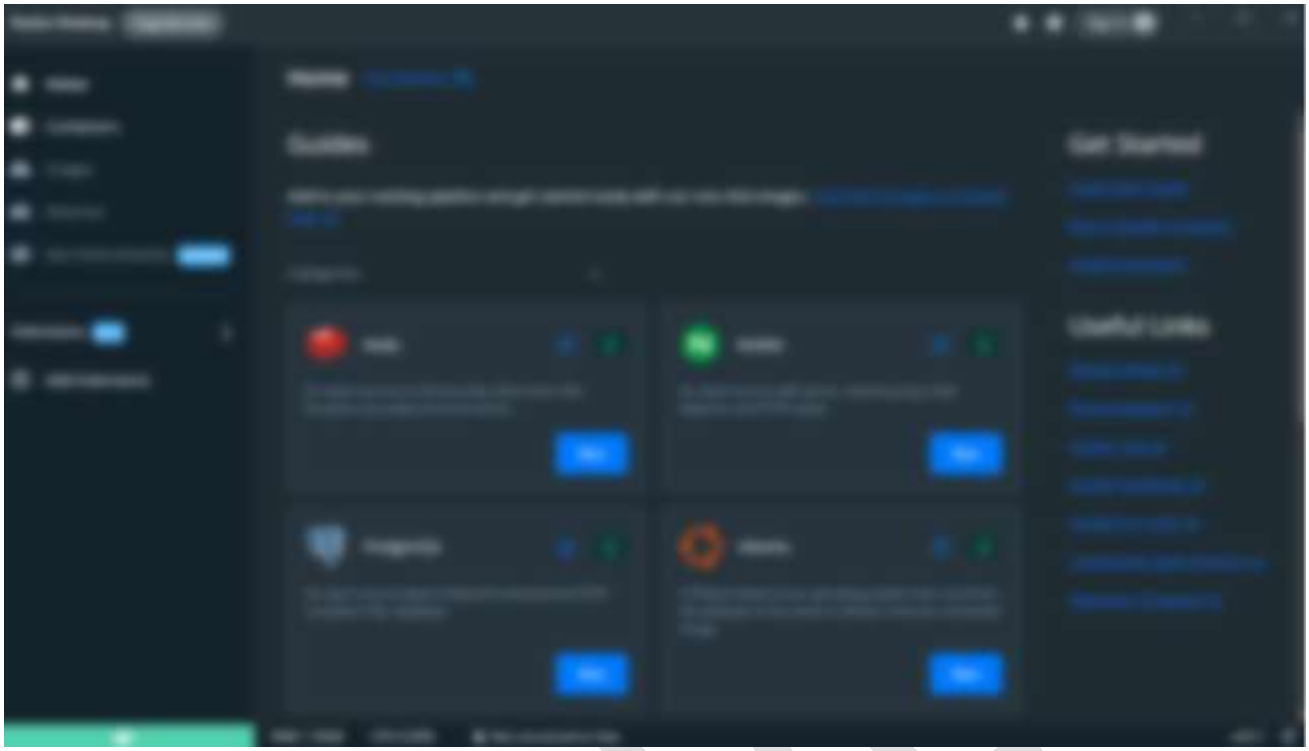


Appendix B- Apache Kafka and Flink





Appendix C- Docker



Appendix D- Operational Monitoring

Appendix E- Formatting Issues and Missing Values

KSB Map

Criteria		Demonstrated in Report Section			
		Exec. Summary and Introduction	Scope of the Project	Outcomes and Findings	Rec. and Conclusions
K2	Outlines the stages of the machine learning lifecycle.	Pg 7	Pg 10		
K7, S2	Explains their use of performance metrics when scoping machine learning engineering solutions by translating business needs and technical problems.	Pg 7-8	Pg 10		
S4	Explains how they apply project management methodologies and techniques for the machine learning activities.		Pg 10-11		Pg 30
K3, K6, S21	Describes the risks of deploying new methods and models and how they assess system vulnerabilities, mitigating the threats or risks to assets and data bias.		Pg 10-11	Pg 15	
K8, S3	Explains how they select and engineer techniques to develop the machine learning solution and the processes used to identify variables and features that can impact stability of model performance during testing and when applying changes to existing models in the live environment.	Pg 7-9	Pg 12		

Criteria		Demonstrated in Report Section			
		Exec. Summary and Introduction	Scope of the Project	Outcomes and Findings	Rec. and Conclusions
K9	Explains feature engineering, selection and pre-processing and their importance in effective machine learning.		Pg 10, 12		
K12, S5	Describes deployment approaches for new data pipelines and automated processes and how they create and deploy models to produce machine learning solutions.		Pg 13		
K16	Explains how programming languages, integrated development environments and modern machine learning libraries are used.		Pg 12		
K10, S7	Explains the techniques applied for output model testing and tuning in order to access accuracy, fit, validity and robustness.		Pg 10	Pg 16	
K13	Outlines the data and information security standard, ethical practices, policies and procedures relevant to data management.		Pg 12-14	Pg 19	
S10, S13, S14	Explains how they track, test and monitor continual learning models into the live environment, ensuring that they remain fit for purpose and stable.		Pg 11-14	Pg 18	
S16	Explains how they transition prototypes into the live environment.		Pg 13		
S8	How they assess system vulnerabilities and mitigate the threats or risks to assets, data and cyber security.		Pg 14	Pg 19	

Criteria		Demonstrated in Report Section			
		Exec. Summary and Introduction	Scope of the Project	Outcomes and Findings	Rec. and Conclusions
K31, S20	Explains how they support the evaluation and validation of machine learning models and statistical evidence to minimise algorithmic bias being introduced using their knowledge of software development best practices.			Pg 23	
K22, S29	Explains how they create and use reports, presentations and other documentation on the model development to confirm stakeholder and end user engagement and approval for handover implementation.			Pg 18, 24	
K23, S28	Explains how they produce and maintain technical documentation using their knowledge of machine learning and data science techniques to enhance the work of other members of the team and meet technical and non-technical user requirements.				Pg 26-28
K17, S15, S23, B2	Explains how they take responsibility for identifying, advocating and embedding ML solutions to deliver environmental and operational sustainable outcomes by considering principles that support organisational strategies and objectives for environmental sustainability.	Pg 7		Pg 21	
K7, S2	Distinction: Evaluates the impact of performance metrics on scoping machine learning engineering solutions when meeting business needs and technical problems.				Pg 25-30

Criteria		Demonstrated in Report Section			
		Exec. Summary and Introduction	Scope of the Project	Outcomes and Findings	Rec. and Conclusions
K10, S7	Distinction: Critically evaluates how their application of techniques ensures the validity and robustness of machine learning.				Pg 25-30
S16, S20	Distinction: Justify how validation of machine learning models minimised bias and the impact this has on prototypes in the live environment.				Pg 25-30


Accelerate People