NeoBank Digital – Analytics & Customer Insights Module

1. Document Control

- **Version:** 1.0
- **Date:** August 2025
- **Author:** Product Team – NeoBank Digital
- **Reviewed By:** Data & Business Intelligence Leads
- **Approved By:** Chief Data Officer

2. Executive Summary

The Analytics & Customer Insights Module empowers NeoBank to leverage data-driven intelligence across lending, repayment, fraud detection, and customer interactions. It enables customer segmentation, predictive analytics, and personalized product recommendations. By consolidating data from multiple modules, this solution helps NeoBank improve customer retention, reduce defaults, and unlock new cross-sell opportunities.

3. Objectives

- Provide real-time dashboards for business stakeholders.
- Enable customer segmentation and behavioral analytics.
- Use machine learning to predict loan defaults and churn.
- Personalize offers and services based on customer insights.
- Support strategic decision-making with advanced analytics.

4. Scope

In-Scope

- Data ingestion from all NeoBank modules.
- Centralized data warehouse & ETL pipelines.
- Customer segmentation and clustering models.
- Predictive models (churn, credit risk, repayment).
- Self-service dashboards and visualization tools.
- Real-time alerts for business KPIs.

Out-of-Scope

- Manual data entry/reporting (covered in Compliance Module).
- Standalone AI/ML platform development.
- Data sourced outside NeoBank ecosystem.

5. Key Features

1. **Unified Data Lake** – store customer, transaction, and loan data.
2. **Segmentation Engine** – cluster customers by demographics, behavior, and risk profile.
3. **Predictive Analytics** – default prediction, churn analysis, cross-sell scoring.
4. **Personalization Engine** – recommend loans, offers, and credit limits.
5. **Interactive Dashboards** – real-time KPIs (loan performance, customer growth, risk).
6. **Alerting System** – threshold-based notifications.
7. **Data Governance Layer** – access controls, lineage, quality checks.

6. User Stories

- As a **marketing manager**, I want customer segmentation so that I can design targeted campaigns.
- As a **risk officer**, I want predictive models so that I can proactively manage defaults.
- As a **relationship manager**, I want personalized recommendations so I can upsell effectively.
- As an **executive**, I want dashboards so that I can monitor KPIs in real time.

7. Integrations

- **Data Warehouse / Data Lake** – central repository (Snowflake, BigQuery, or AWS Redshift).
- **ETL Pipelines** – Apache Airflow, dbt, or Spark.
- **Visualization Tools** – Power BI, Tableau, or Looker.
- **ML Frameworks** – TensorFlow, PyTorch, scikit-learn.
- **Upstream Modules** – Lending, Repayment, Fraud, Compliance, API Ecosystem.

8. Authentication & Security

- Role-based access control for dashboards & reports.
- Data masking for sensitive attributes.
- End-to-end encryption (AES-256 + TLS 1.3).
- Audit logs of queries and report access.
- Compliance with GDPR and local data privacy laws.

9. Non-Functional Requirements (NFRs)

- **Performance:** Dashboards refresh in ≤ 5 seconds.
- **Availability:** 99.9% uptime for analytics services.
- **Scalability:** Handle 10M+ customer records.
- **Compliance:** GDPR, CCPA, RBI data guidelines.
- **Security:** Data anonymization for analytics workloads.

10. Risks & Mitigations

Risk	Impact	Mitigation
Poor data quality	Misleading insights	Data governance and validation pipelines
High infrastructure costs	Budget overruns	Cloud cost optimization and monitoring
ML model inaccuracies	Wrong predictions	Continuous retraining and monitoring
Unauthorized access	Data breaches	RBAC, encryption, audit logs

11. Success Metrics

- Dashboard Adoption: ≥ 80% of target users.
- Churn Prediction Accuracy: ≥ 85%.
- Default Prediction Accuracy: ≥ 90%.
- Increase in Cross-Sell Conversions: ≥ 15%.
- Reduction in Manual Reporting Effort: ≥ 70%.

12. Dependencies

- Centralized data infrastructure availability.
- Upstream data quality from other modules.

- ML engineering resources for predictive models.
- BI tool licenses.

13. Timeline & Milestones

- **Week 1–2:** Requirements & architecture finalization.
- **Week 3–4:** Data lake & ETL pipeline setup.
- **Week 5–6:** Customer segmentation engine.
- **Week 7–8:** Predictive model development.
- **Week 9:** Dashboard & visualization integration.
- **Week 10:** Real-time alerting system.
- **Week 11:** UAT & stakeholder review.
- **Week 12:** Production rollout.

14. Appendix

- Glossary: ETL, KPI, Churn, Default, Cross-Sell.
- References: GDPR, CCPA, RBI Analytics Guidelines.