Research Report on Data Analytics, Data Science, and AI in Customer Churn Prediction

This report provides a theoretical foundation for churn prediction—identifying customers at risk of leaving a service—particularly in telecommunications datasets.

Define Key Terms

Data Analytics: Describes and diagnoses churn patterns using statistics and visuals.

Data Science: Designs end■to■end pipelines—ingestion, cleaning, feature engineering, modeling, and evaluation.

Artificial Intelligence (AI): Learns from labeled historical churn to predict risk and recommend retention actions.

Business Impact of Churn

Churn directly affects recurring revenue and acquisition costs. Predicting churn enables targeted retention offers, service quality improvements, and proactive outreach.

Literature & Common Approaches

Frequent algorithms include Logistic Regression, Decision Trees, Random Forests, Gradient Boosting, and SVMs. Typical challenges: class imbalance, leakage, high■cardinality categorical features, and dynamic behavior signals (tenure, contract, charges).

Telecom Domain Considerations

Key drivers often include contract type (month■to■month vs. term), tenure, monthly/total charges, payment method, add■on services, and service quality. Regulatory compliance and portability rules may influence churn behavior.

Ethical Considerations

Protect PII, avoid unfair discrimination in retention offers, audit models for bias/drift, and ensure transparency of automated decisions that affect customers.

Conclusion

Churn prediction aligns analytics, data science, and AI to reduce revenue loss and improve customer experience via timely, explainable interventions.