Auto Insurance Churn Detection

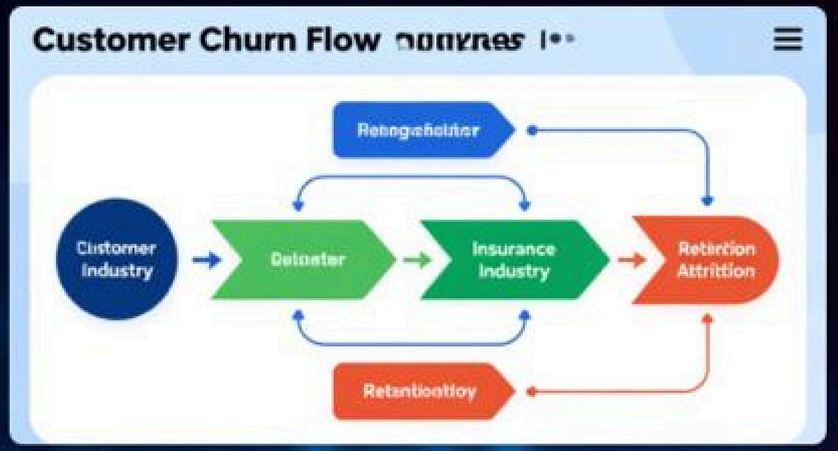
Advanced ML Model with Explainable AI

by Team rm -rf

Project Overview & Challenges

Business Context

Auto insurance churn costs companies millions in lost revenue. Early identification of at-risk customers enables proactive retention strategies, reducing acquisition costs and maintaining market share.



High Churn Impact

Industry average 15-25% annual churn rate translates to significant revenue loss and increased acquisition costs.

Prediction Complexity

Multiple factors influence churn: pricing, claims history, customer service, competitor offerings, and life circumstances.

Solution

- Use machine learning models to predict customer churn probabilities.
- Apply AI explainability metrics like SHAP to understand which features most influence churn predictions.
- Target high-risk customers with personalized retention strategies based on these insights.

Advantages:

- Improves retention by acting on data-driven predictions.
- Enhances trust with transparent, explainable AI insights.
- Optimizes marketing spend by focusing on customers most likely to churn.

Model Architecture & Training

XGBoost Advantages

Superior handling of tabular data, automatic feature selection, robust to outliers, and excellent predictive performance with interpretable results.

S

Neural Network Challenges

Risk of overfitting with complex architectures, requires careful regularization, and needs substantial computational resources for training.

W

Regularization Threats

Over-regularization can lead to underfitting, requires hyperparameter tuning, and may increase training time significantly.

T

MLP Opportunities

Ability to capture complex non-linear relationships, flexibility in architecture design, and potential for ensemble integration.

O

Multi-Model Confidence Scoring

XGBoost Predictions

Gradient boosting model provides robust churn probability estimates with feature importance rankings for interpretable results.

Deep Neural Network

Multi-layer perceptron captures complex non-linear relationships and interactions between customer features and churn behavior.

Ensemble Confidence

Weighted ensemble combines both models with confidence scoring based on prediction agreement and historical accuracy metrics.

Confidence score

$$\text{Confidence} = f(\text{Agreement}, \text{Certainty})$$

Explainable AI with SHAP

SHAP Value Calculation

Computed Shapley values for each prediction, quantifying individual feature contributions to churn probability outcomes.

Feature Importance Analysis

Identified key drivers: claim frequency, premium changes, customer tenure, and service interactions as primary churn indicators.

Business Interpretation

Translated technical insights into actionable business strategies for customer retention and personalized intervention programs.

Model Performance & UI Integration

91%
Prediction Accuracy

0.90
AUC-ROC Score

>85%
Confidence Score

Clean UI Design

> Intuitive dashboard interface enables business users to easily interpret predictions and take proactive retention actions.

- Real-time churn probability display
- Customer risk segmentation visualization
- Actionable recommendation engine

AI-Powered Insights

> Automated insight generation provides context-aware recommendations for customer retention strategies.

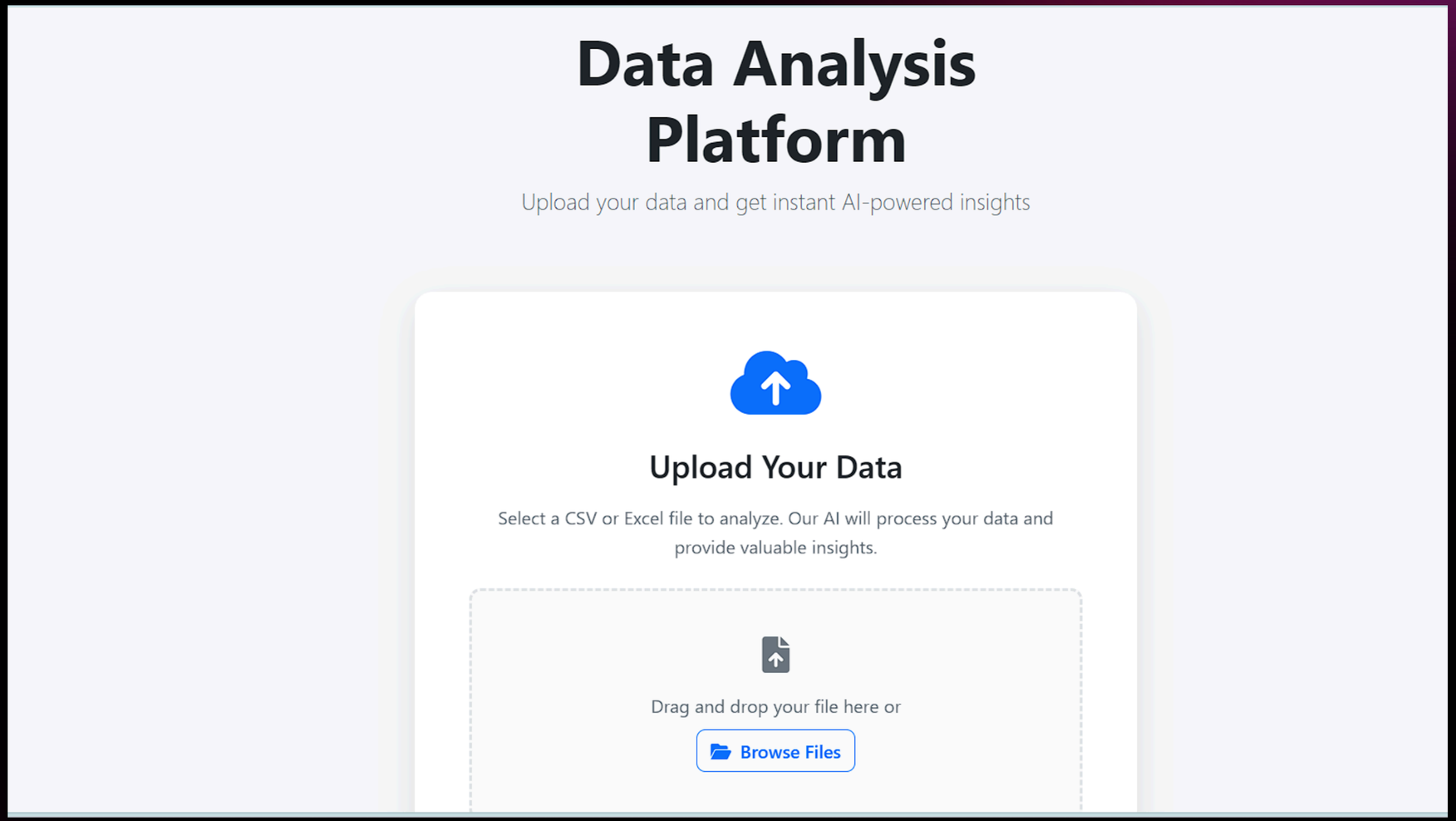
- Personalized intervention suggestions
- Predictive customer lifetime value
- Next-best-action recommendations

Ensemble Robustness

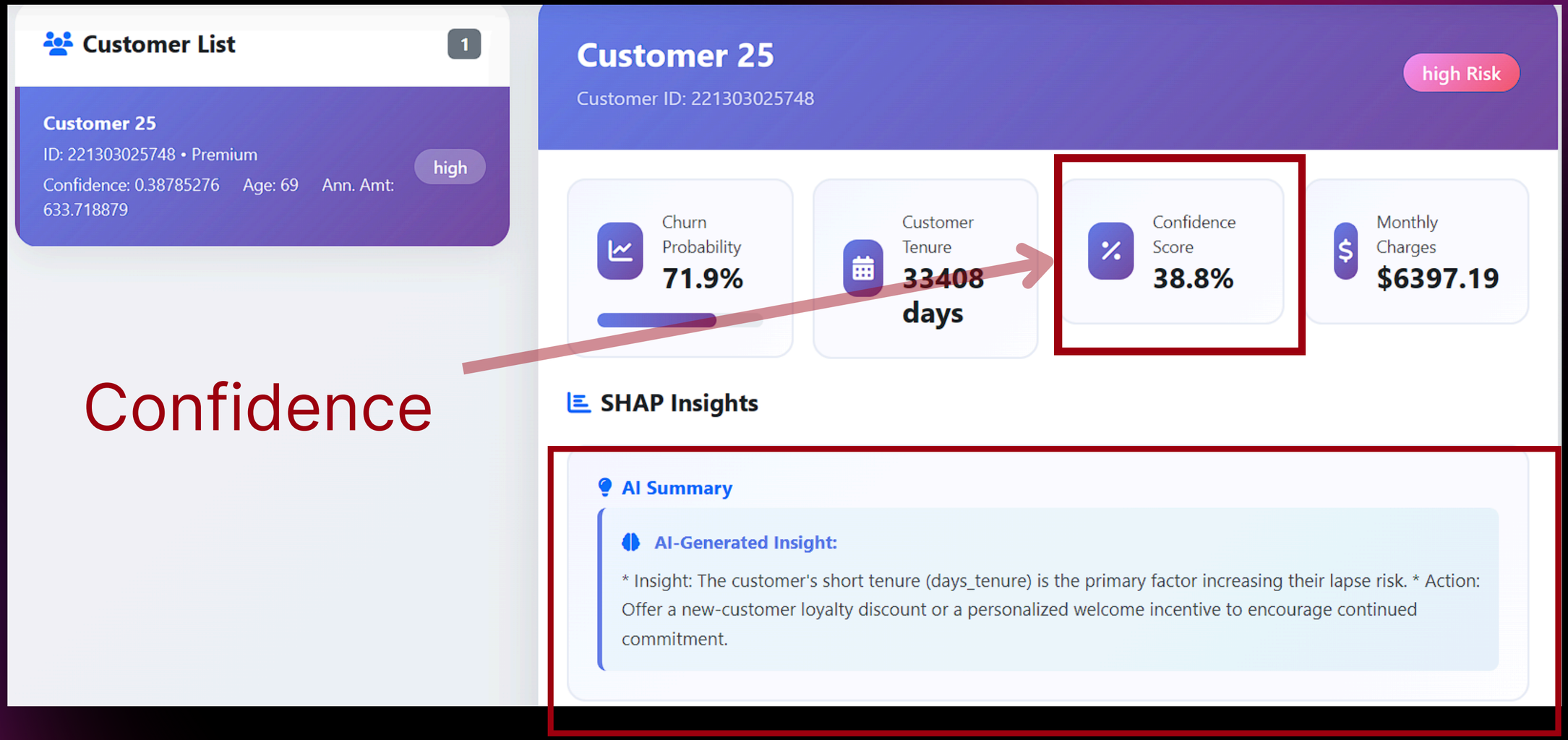
> Multiple model ensemble ensures reliable predictions across diverse customer segments and market conditions.

- XGBoost + MLP weighted ensemble
- Confidence scoring mechanism
- Model drift detection and adaptation

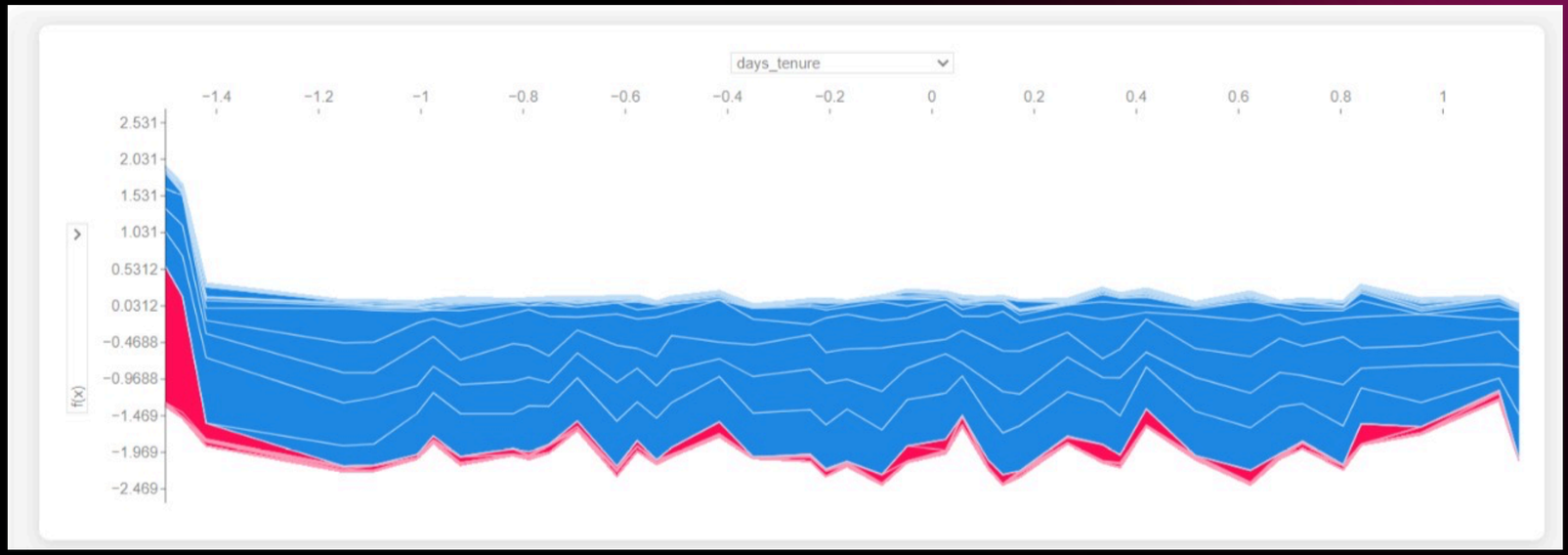
Dump your data !



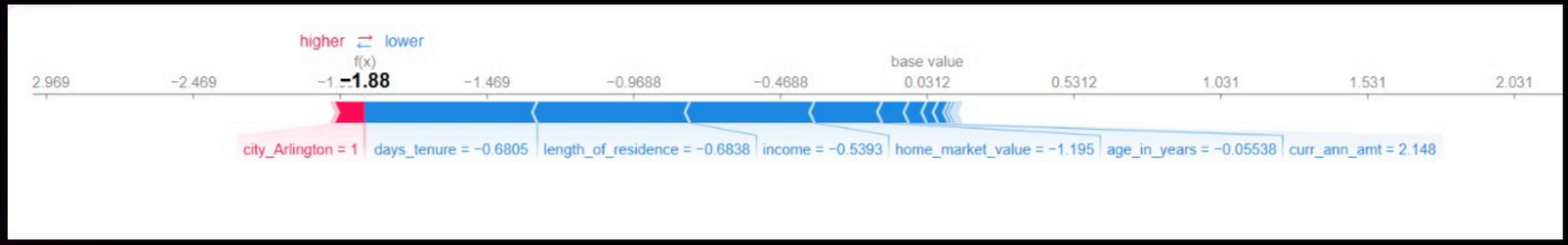
get insights and recommendations from AI !



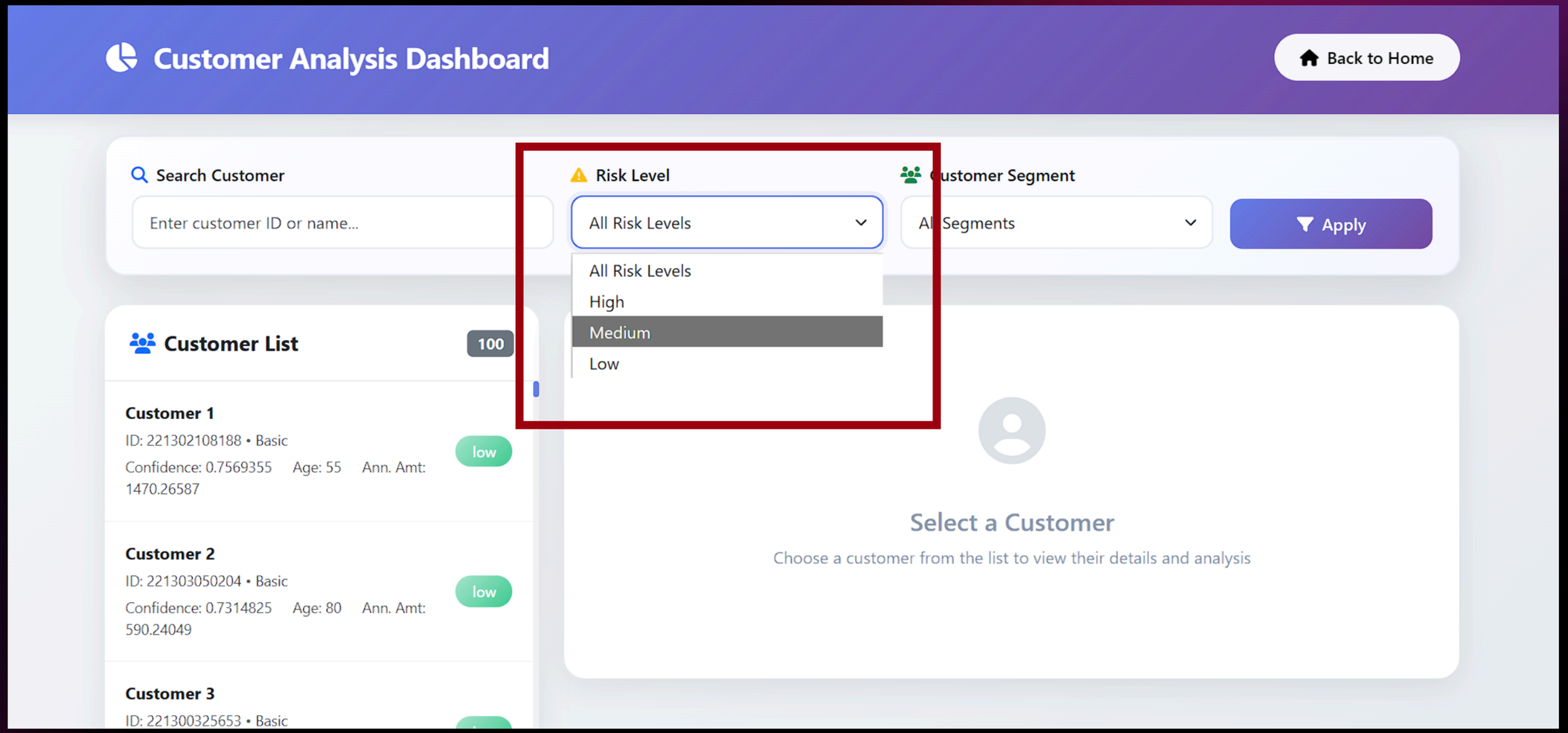
visualize the whole data



visualize per user



Risk classification



Future scope

- Future plans for model expansion:
- Incorporate regional data to capture geographic patterns.
- Extend predictions to motorcycle and other insurance lines.
- Analyze trends to identify high-risk customer segments, e.g., new customers, high-premium policies, or regions with historically higher churn rates.

Thank you !