Transfer IQ: AI-Driven Player Valuation

Predicting Market Value using Deep Learning and Multi-Source Data

Presented by: P.VEERABABU I October 9, 2025

Data Collection & Sourcing

Aggregating comprehensive data streams for holistic player assessment.

Performance Metrics

StatsBomb Open Data (Event-level metrics), FBref, Futbin.

Market Data

Transfermarkt: Historical value trends and contract status.

Public Sentiment

Twitter/Reddit: VADER NLP for crucial sentiment scoring.

Fitness & Risk

Injury history translated into a total_days_missed risk feature.

Data Cleaning & Preprocessing

Establishing a robust foundation for accurate modeling.

01

Data Quality Checks

Handling missing values, outliers, and inconsistent formats for reliable input.

02

Feature Scaling

Normalize and standardize numerical features for model compatibility.

03

Encoding Variables

Converting categorical data (e.g., position, nationality) into machine-readable formats.

Feature Engineering

Creating high-impact predictors beyond raw statistics.

Advanced Metrics

Deriving form trends, injury risk scores, and growth potential indicators.

Correlation Analysis

Selecting the most impactful features that significantly influence market value.

External Context

Incorporating transfer market trends, contract length, and current club status.

The output is the **master_list_cleaned.csv**—the final, optimized dataset ready for machine learning.

Model Development I - Baselines

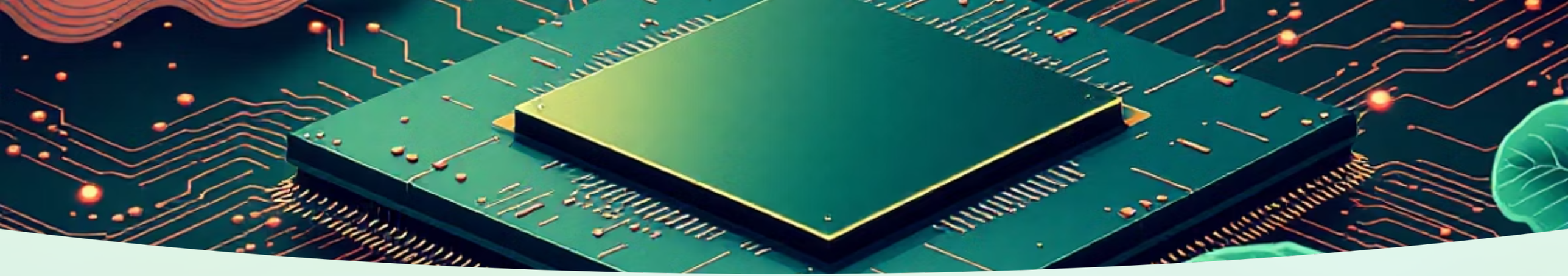
Initial comparative analysis using established regression techniques.

Baseline Models

Training Multiple Linear Regression and Random Forest Regression.

Performance Evaluation

Measuring predictive power using MAE, RMSE (error), and R^2 scores (fit).



Model Development II - Deep Learning

Adopting XGBoost for superior prediction accuracy.

XGBoost Selection

Chosen for its speed, performance, and robustness in handling complex, non-linear data structures.

Non-Linear Relationships

Excels at capturing subtle interactions between diverse feature sets (performance, market, sentiment, fitness).

Regularization

Built-in mechanisms to reduce the risk of overfitting and improve generalization on unseen players.

Evaluation & Tuning

Ensuring the model is robust, accurate, and ready for real-world application.

Cross-Validation

Rigorous testing to ensure model stability and reliability across different data subsets.

Hyperparameter Tuning

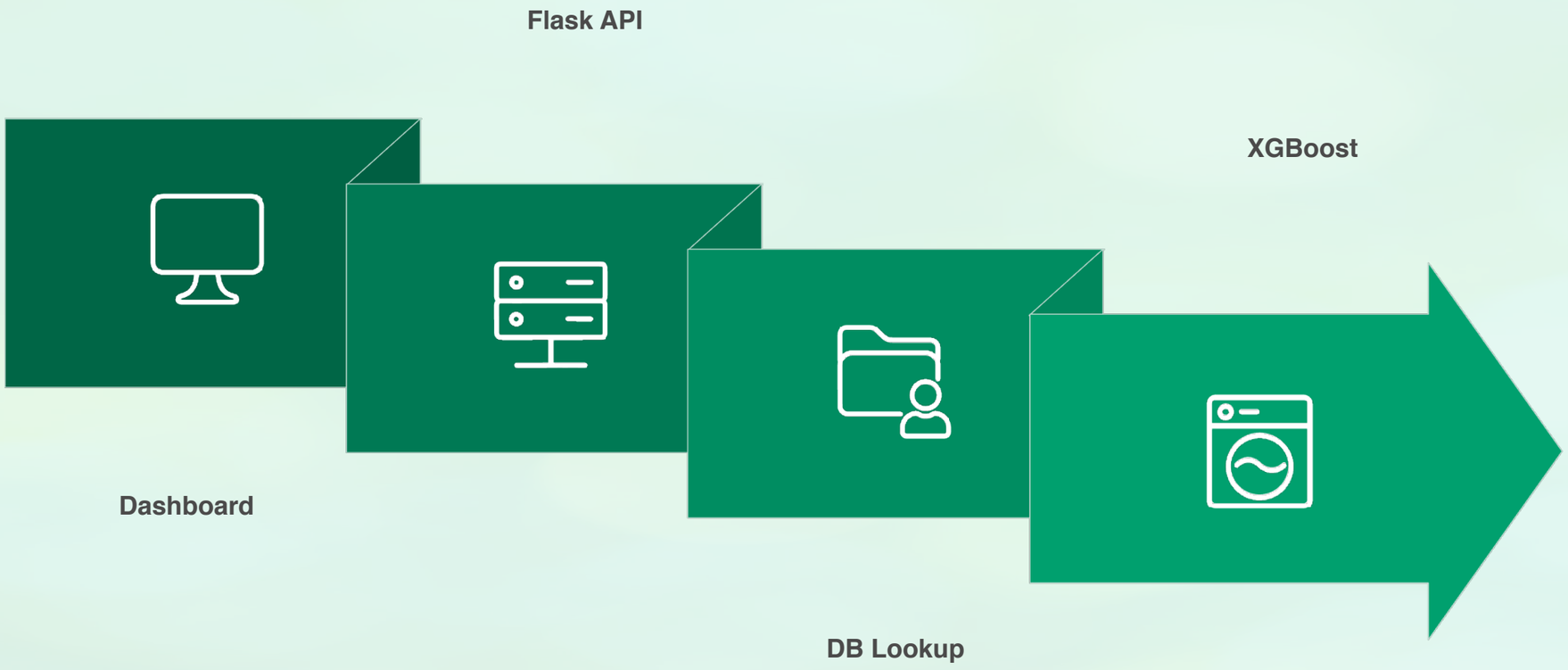
Fine-tuning XGBoost parameters to maximize predictive accuracy (optimizing MAE/RMSE).

Generalization Check

Validation against unseen test data to confirm performance in real-world scenarios.

Deployment Architecture

Delivering instant, actionable valuation insights.



Transfer IQ: Conclusion

Enhanced Efficiency

AI-driven valuation sharpens transfer market insight and reduces financial risk.

Strategic Edge

Empowers clubs to identify undervalued talent and negotiate smarter deals.

Objective Valuation

Represents the future of transparent and objective player assessment.

Thank You & Q&A