BASKETBALL TOURNAMENT PREDICTION SYSTEM:

WE ALREADY KNOW THE OUTCOME—WHY JUST PLAY?

By:
- Nicolas Romero Rodriguez
- Carlos Andres Celis Herrera
- Xiomara Salome Arias Arias

WHY A BASKETBALL
TOURNAMENT SYSTEM?

- Kaggle competition.
- Based in college basketball tournament organized by the NCAA
- Direct elimination, 68 teams and 7 rounds
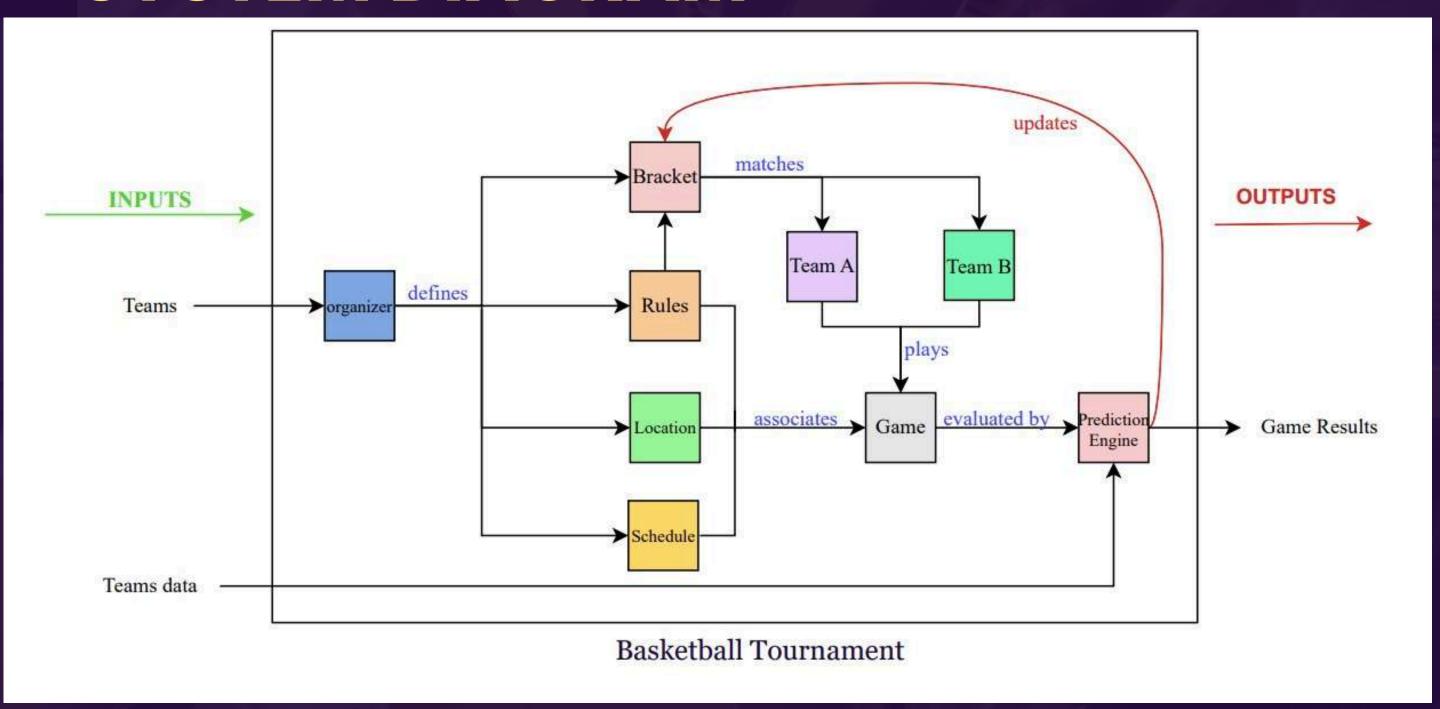
HOW TO PREDICT
THE MATCHES?

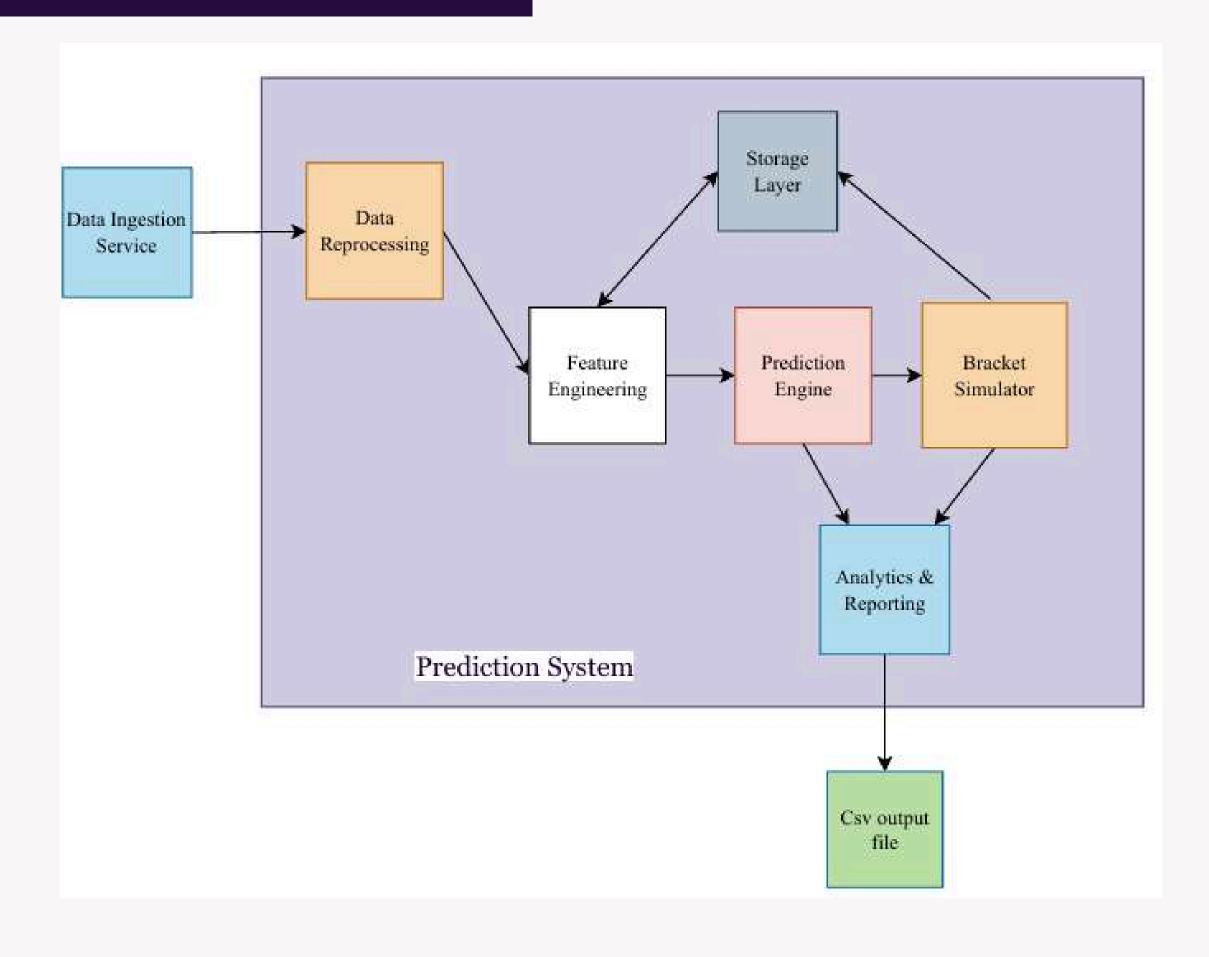- *Design and implement an autonomous prediction system capable of forecasting the results of tournament games.*
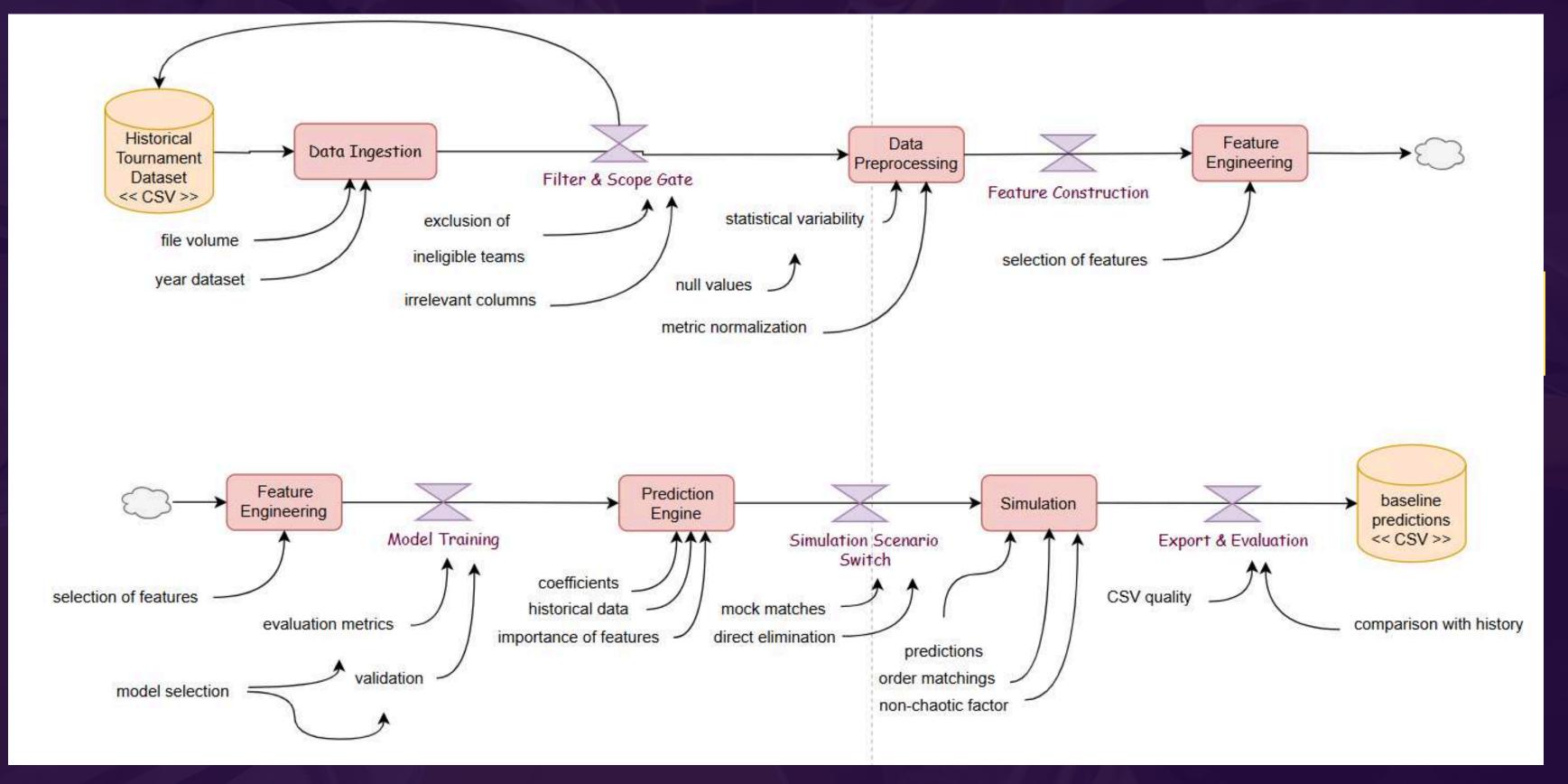
ARCHITECTURE DIAGRAM

BASKETBALL TOURNAMENT
STOCK AND FLOW

SYSTEM ARCHITECTURE

CHARACTERISTICS

✅ **Environmental factors**

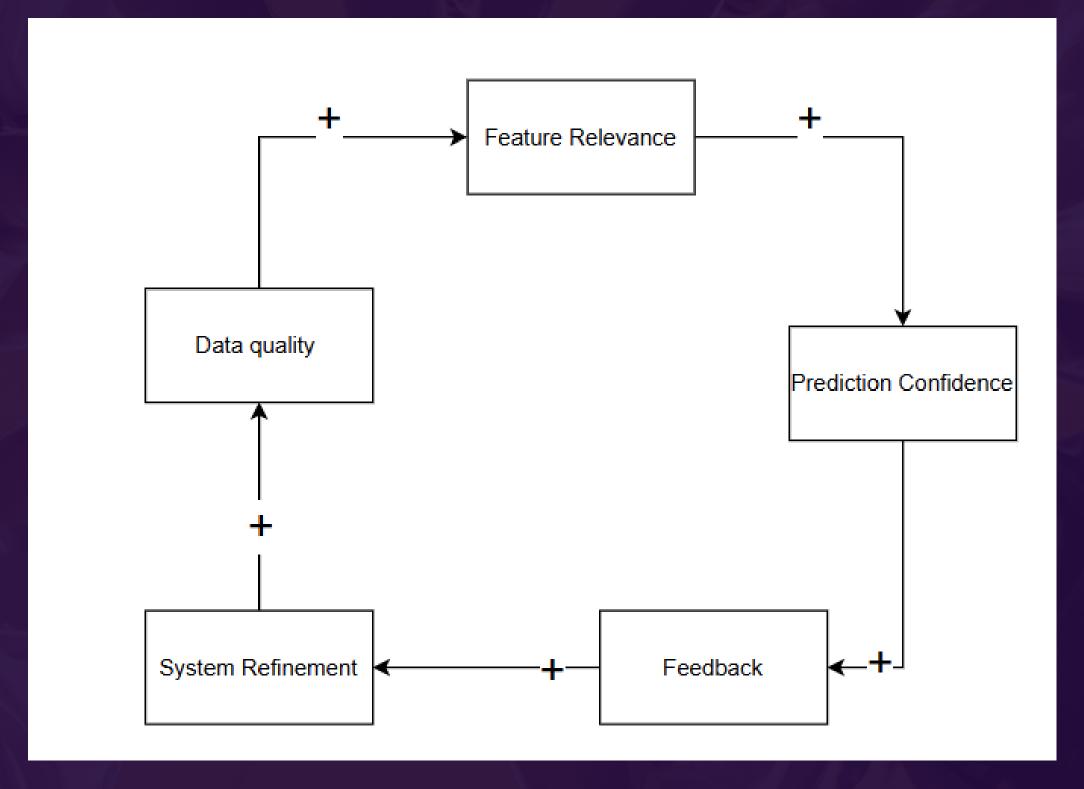*It's not possible to measure environmental impacts and random factors*

✅ **Modularity and independence**

Each module of this architecture is assigned a single responsability to ensure modularity.

✅ **Autonomy**

From raw data ingestion to CSV output, the system automates preprocessing, matchup generation, feature extraction, prediction, and result formatting
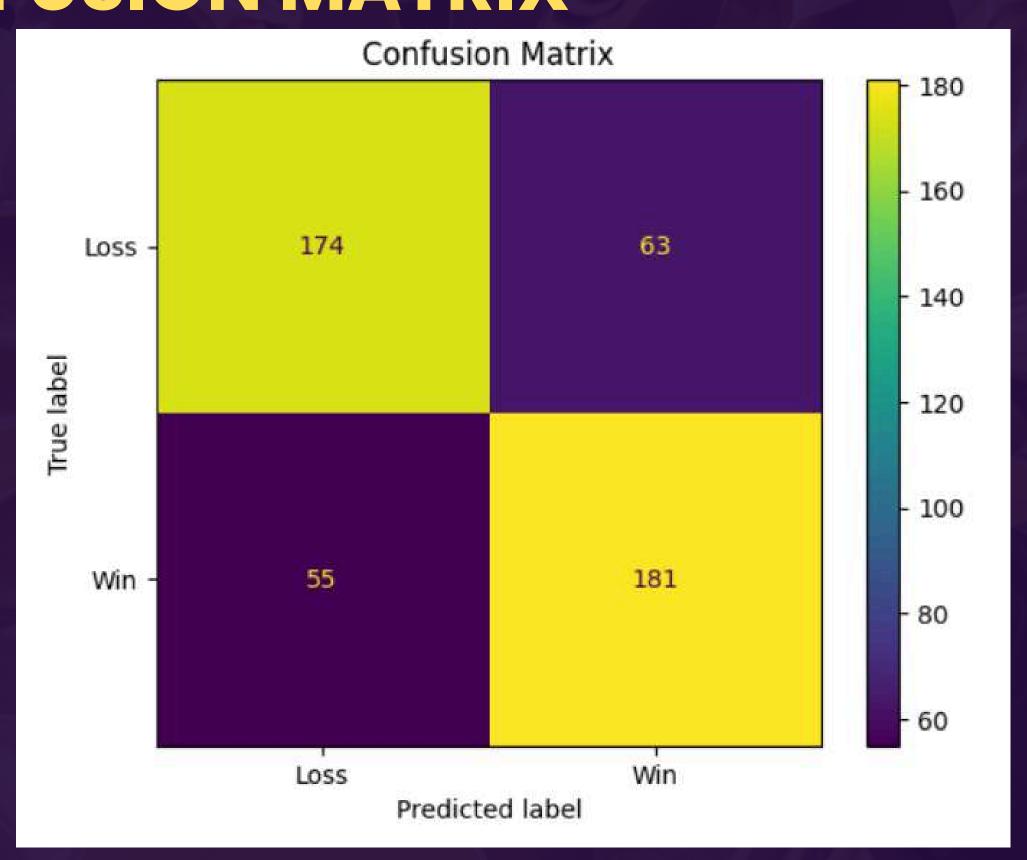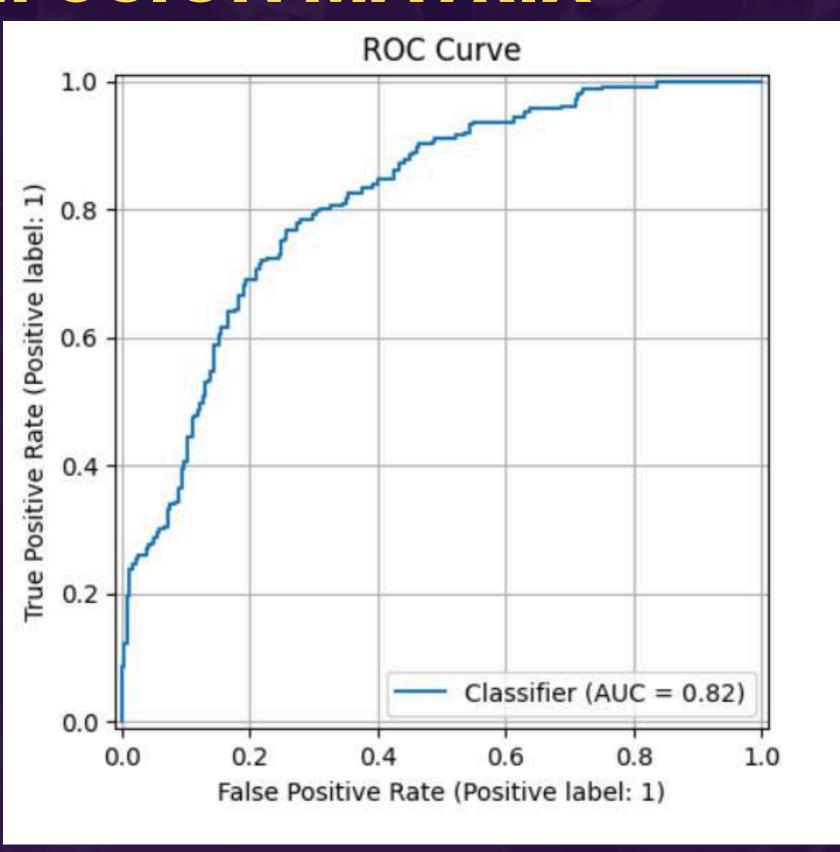
FEATURES

Season	Gender	TeamA	TeamB	SeedDiff	WinLossDiff	MarginDiff	PPGDiff	PtsAllowedDiff	Winner
2015	M	1214	1264	0	-0.108902	-3.223485	-2.723485	0.500000	1
2015	M	1279	1140	0	-0.093750	-3.781250	-9.218750	-5.437500	1
2015	M	1173	1129	0	0.015640	-1.632454	-1.658847	-0.026393	1
2015	M	1352	1316	0	-0.069404	-4.146628	-4.935484	-0.788856	1
2015	M	1112	1411	13	0.264706	17.088235	8.264706	-8.823529	1

BASKETBALL TOURNAMENT
CONFUSION MATRIX
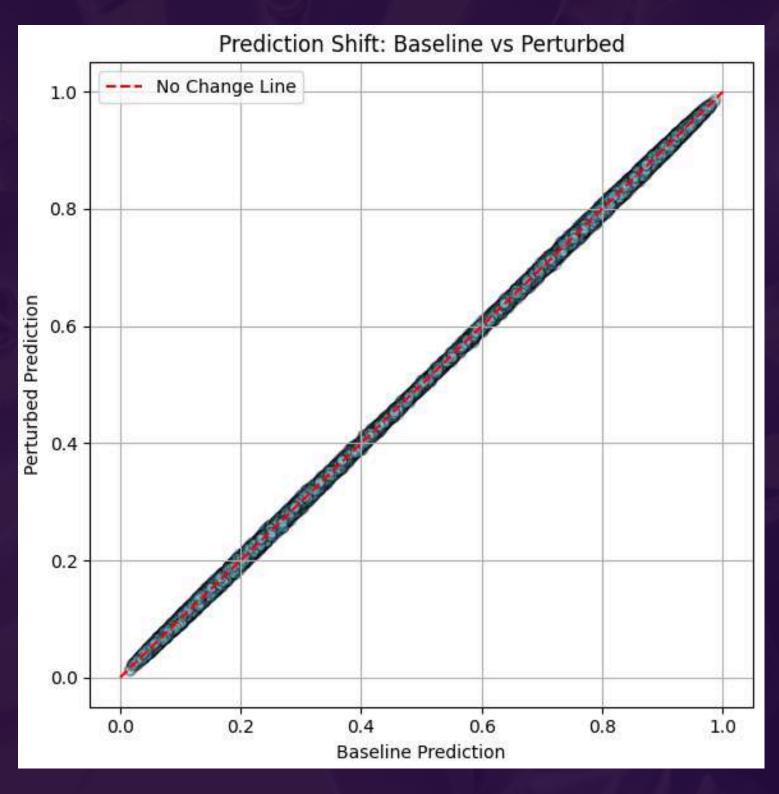
Confusion Matrix
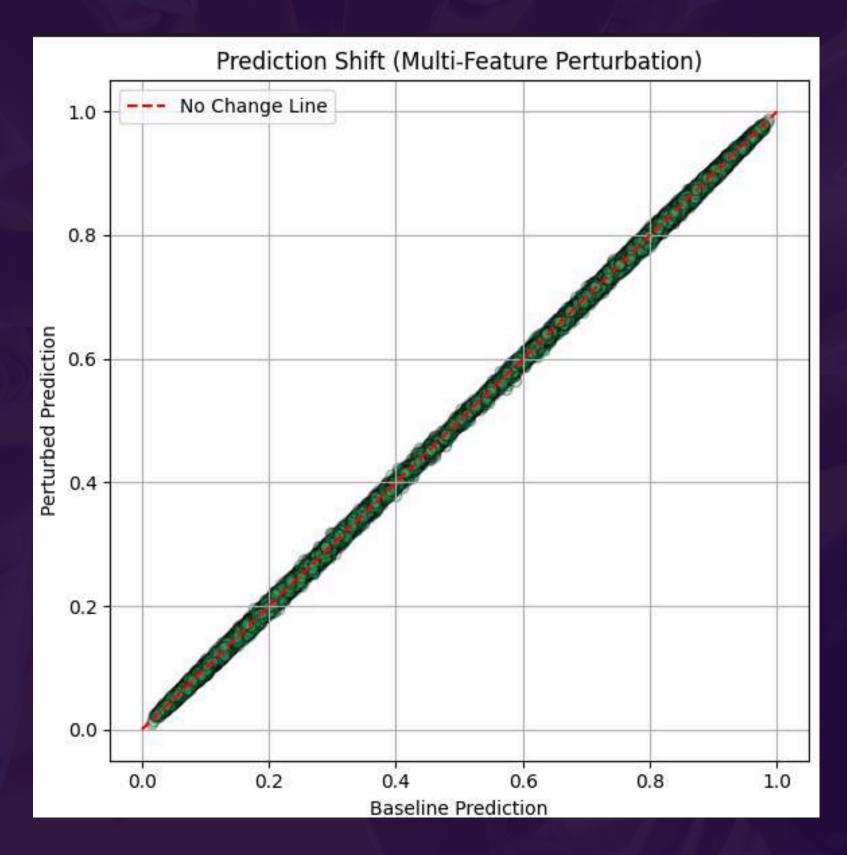
BASKETBALL TOURNAMENT
CONFUSION MATRIX

PERTURBED SCENARIO: SENSITIVITY TO INPUT NOISE

Prediction Shift: Baseline vs Perturbed
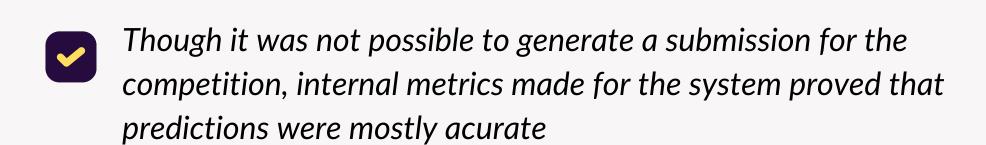
PERTURBED MULTI FEATURE SCENARIO

Prediction Shift (Multi-Feature Perturbation)

CON<mark>CLUSIONS</mark>

✅ *Though it was not possible to generate a submission for the competition, internal metrics made for the system proved that predictions were mostly acurate*

✅ *Adding additional noise to the dataset doesn't seem to affect the prediction quality*

THANKS