



FOOTBALL DATA ANALYTICS DECISION-SUPPORT SYSTEM

A transparent, selective and chronological pre-match analysis pipeline

ALEXANDRU CARP | BSc Applied Computing | Student ID 2317094

ACCA6020
Final Project Portfolio
10 August 2026

RESEARCH QUESTION

How can historical football data be combined to produce explainable, high-confidence decisions without forcing a selection?

01 PROBLEM AND AIM

Football prediction is often reduced to averages, while line-ups, injuries, tactical changes and contradictory evidence are overlooked. Some systems also force a prediction even when the evidence is weak.

The aim was to build an explainable decision-support system that combines multiple evidence sources and can deliberately return NO_ELITE_PICK.

02 PROJECT OBJECTIVES

1. Review sports analytics and selective prediction methods [1,2].
2. Build a structured multi-season football database.
3. Combine team, player, line-up, injury, formation and context evidence.
4. Apply strict SELECT / NO_ELITE_PICK rules.
5. Evaluate accuracy, selectivity, chronology and technical reliability.

03 DATA AND SCOPE

SQLite database populated from API-Football

7,365 completed fixtures across 2016-2026

Three Scandinavian competitions:

- Eliteserien
- Superettan
- Allsvenskan

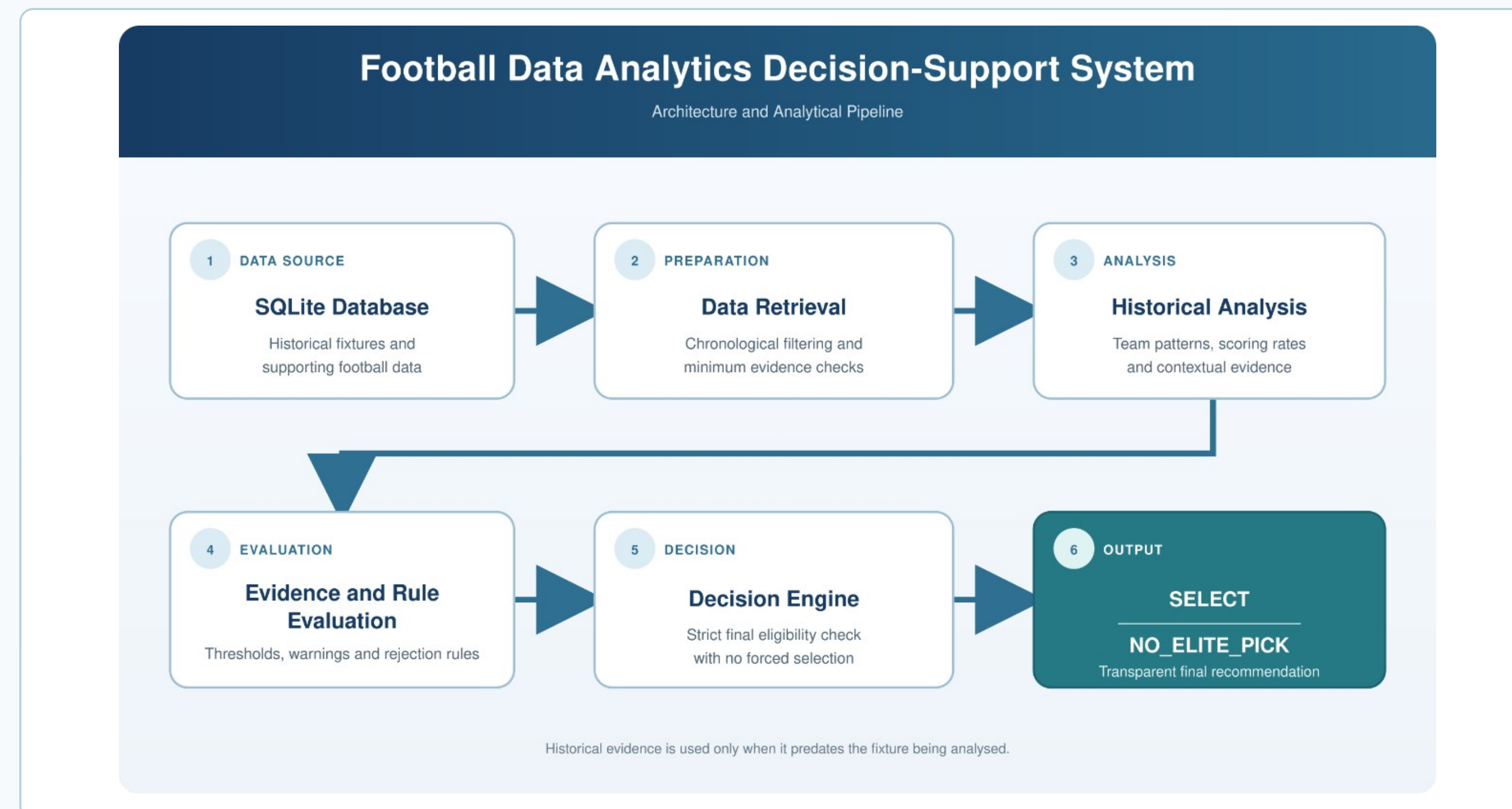
Primary decision market: Match Over 1.5 full-time goals

CHRONOLOGICAL RULE
Only evidence recorded before kick-off can influence a decision.

CONCLUSION

The project demonstrates a working, transparent decision-support system that prioritises disciplined abstention over high prediction volume.

04 SYSTEM ARCHITECTURE



05 IMPLEMENTATION EVIDENCE



KEY DESIGN PRINCIPLE

“More data should make the system stricter, not more active.”

Weak, missing or contradictory evidence reduces confidence or triggers NO_ELITE_PICK - a selective reject-option approach [2].

06 EVALUATION RESULTS

7,365
historical fixtures analysed

118
fixtures selected

Accepted cases correct **91.53%**

Fixtures rejected **98.40%**

108 correct / 10 incorrect | 7,247 NO_ELITE_PICK decisions
Source: Author's own system evaluation

17 / 17 correct selections in the 2025 holdout season (100%, small sample)

13 / 13 automated technical tests passed
Leakage audit: PASS

07 LIMITATIONS AND FUTURE WORK

LIMITATIONS

- Historical validation does not prove live performance.
- Coverage 1.60%; 2025 holdout: 17 selections.
- Historical odds/context records are incomplete.

FUTURE WORK

- Prospective logging, odds/value assessment, wider leagues, calibrated confidence and stronger line-up context.

REFERENCES

- [1] Bunker & Susnjak (2022), JAIR, 73, 1285-1322.
[2] Geifman & El-Yaniv (2017), NeurIPS, 30.