



Football tracking data를 이용한 이벤트 데이터 예측 및 분석

팀명: fm마스터

부산대학교 정보컴퓨터공학부

202055614 최성민

202055502 강동권

지도교수: 이기준

목차

- 1 과제 배경 및 목표
 - 1.1 과제 배경
 - 1.2 과제 목표
- 2 요구 사항 분석
 - 2.1 기능적 요구사항 분석
 - 2.2 비기능적 요구사항 분석
- 3 개발 환경 및 사용 기술
 - 3.1 개발 환경
 - 3.2 사용 기술 및 프레임 워크
 - 3.3 외부 리소스 및 참고 데이터
 - 3.4 Workflow Diagram
- 4 현실적 제약 사항 분석 결과 및 대책
- 5 개발 일정 및 역할 분담
 - 5.1 개발 일정
 - 5.2 역할 분담

1. 과제 배경 및 목표

1.1 과제 배경

현대 스포츠에서는 데이터 기반의 전략 수립과 경기 분석이 점점 더 중요해지고 있으며, 그 중심에는 트래킹 데이터(tracking data)가 있다. 트래킹 데이터란 경기 중 선수들의 위치, 이동 경로, 속도, 가속도 등 시공간적인 정보를 일정 주기로 수집한 데이터로, 전통적인 이벤트 기반 데이터(예: 패스, 슈팅 기록 등)보다 훨씬 더 정밀하고 방대한 정보를 담고 있다. 이러한 데이터는 선수의 움직임 패턴, 포지셔닝 전략, 팀 전술 분석 등 다양한 목적으로 활용될 수 있다.

특히 축구의 경우 22명의 선수와 공이 넓은 필드 위를 끊임없이 움직이며 복잡한 상호작용을 보이기 때문에, 경기 이해를 위해 트래킹 데이터를 활용한 분석이 필수적으로 여겨지고 있다. 하지만 현재 대부분의 분석 시스템은 사후적이고 정적이며, 실제 경기 상황에서 일어나는 이벤트를 실시간으로 예측하거나 맥락적으로 해석하는 데에는 한계가 있다.

예를 들어, 패스나 슈팅과 같은 이벤트는 단순히 공의 움직임만으로 설명되기 어렵고, 주변 선수들의 위치, 압박 유무, 속도 변화 등 다양한 요소가 결합되어 발생한다. 따라서 트래킹 데이터를 기반으로 한 이벤트 예측 모델은 보다 높은 정확도와 풍부한 경기 해석을 가능하게 하며, 나아가 코칭 스태프나 전술 분석가에게 실질적인 인사이트를 제공할 수 있다.

또한, 머신러닝과 딥러닝 기술의 발전으로 이러한 예측 문제를 해결할 수 있는 가능성이 더욱 커지고 있다. 시계열 분석, 공간적 관계 모델링, 행동 패턴 인식 등의 기법을 트래킹 데이터에 적용하면, 경기 중 발생 가능한 주요 이벤트를 사전에 예측하고, 이를 기반으로 전략적인 결정을 내릴 수 있다.

따라서 본 과제는 축구 트래킹 데이터를 활용하여 이벤트 발생을 예측하고 이를 분석하는 시스템을 개발함으로써, 기존 분석의 한계를 극복하고 보다 지능적인 스포츠 분석 플랫폼의 기틀을 마련하는 데 의의가 있다.

1.2 과제 목표

본 과제는 축구 경기에서 수집된 선수 tracking data를 기반으로 다음의 세 가지 목표를 달성하는 것을 목적으로 한다.

- **이벤트 예측 모델 개발:** 선수들의 위치와 움직임 데이터를 기반으로 패스, 슈팅, 태클 등 주요 이벤트를 예측할 수 있는 모델을 설계 및 구현한다. 여러가지 딥러닝 모델을 Survey하여 모델을 선택하고 Tuning한다.
- **Tracking data 가공:** Football Manager 2024 시뮬레이터(이하 FM2024)를 통해서 축구 경기에서의 Tracking data를 여러가지 기법을 통해 직접 추출하고 가공하여서 데이터 부족 문제를 해결한다.
- **이벤트 분석 및 시각화:** 예측된 이벤트 데이터를 활용하여 경기 흐름, 전략, 선수별 기여도 등을 분석하고 이를 직관적으로 시각화할 수 있는 시스템을 구축한다.

2. 요구 사항 분석

2.1 기능적 요구사항 분석

- (1) 트래킹 데이터 입력 기능 (CSV, JSON 등 다양한 포맷 지원)

FM2024 또는 실제 경기로부터 수집된 트래킹 데이터를 다양한 포맷(CSV, JSON 등)으로 입력 받을 수 있어야 한다.

- (2) 영상 기반 트래킹 데이터 추출 기능

FM2024의 경기 영상(예: 화면 녹화)을 기반으로 컴퓨터 비전 기술을 통해 각 선수의 좌표 및 이동 경로를 자동 추출하는 기능이 있어야 한다.(예: OpenCV, YOLO 등 활용)

- (3) 이벤트 예측 기능

추출된 트래킹 데이터를 기반으로 주요 이벤트(패스, 슈팅, 태클 등)를 예측할 수 있는 머신러닝 또는 딥러닝 기반 모델을 사용한다.

- (4) 이벤트 시각화 및 분석 기능

예측된 이벤트를 필드 위에 시각화하고, 시간 흐름에 따른 이벤트 흐름 또는 팀별/선수별 분석 리포트 제공

- (5) 선택적 기능사항

시나리오별 분석기능, 사용자 인터페이스 등

2.2 비기능적 요구사항 분석

- 확장성: 다양한 경기 데이터나 리그에 적용 가능해야 한다.
- 성능: 실제 Ground Truth 값에 근접해야 하며, 유의미한 결과를 도출해야 한다.
- 사용성: 데이터 과학이나 축구 분석에 익숙하지 않은 사용자도 이해할 수 있도록 설계
- 호환성: 다양한 플랫폼(Windows/Linux) 및 포맷과 호환 가능해야 함

3. 개발 환경 및 사용 기술

3.1 개발환경

항목	내용
OS	Windows 11 / Ubuntu 22.04 (개인 개발 환경 기준)
프로그래밍 언어	Python 3.10 / Python 3.9
개발 도구	Jupyter Notebook, Visual Studio Code
버전 관리	Git, GitHub
가상환경	conda 또는 venv

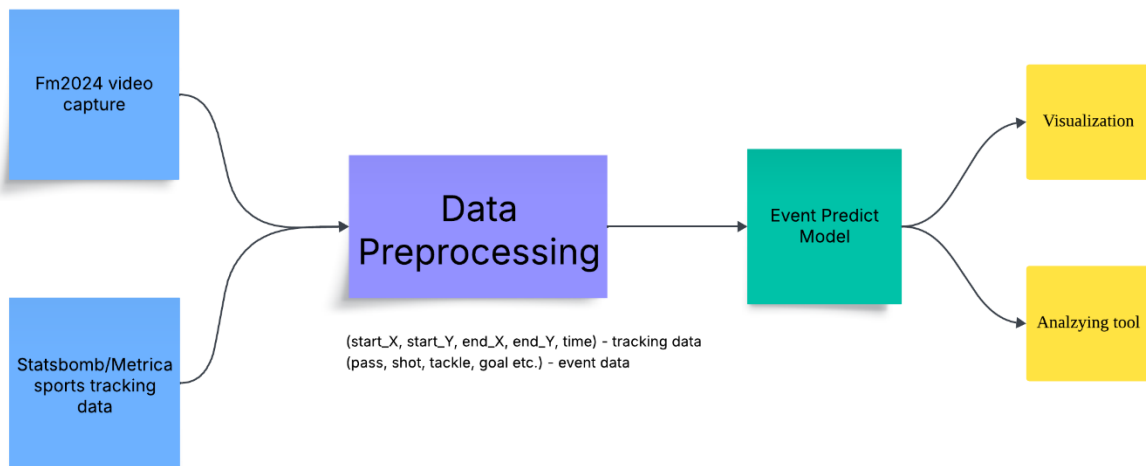
3.2 사용 기술 및 프레임 워크

- (1) 영상 처리 및 프레임 추출: FFmpeg, OpenCV
- (2) 객체 검출 및 트래킹: YOLOv8, Deep SORT, Numpy, pandas
- (3) 좌표 정규화 및 필드 맵핑: Metrica IO, SciPy
- (4) 이벤트 예측 모델링: scikit-learn, PyTorch/TensorFlow, Optuna
- (5) 데이터 시각화 및 분석: Matplotlib, Seaborn, Plotly, Pandas-Profiling
- (6) 모델 평가 및 성능 측정: Classification Metrics, Cross-validation

3.3 외부 리소스 및 참고 데이터

- (1) FM2024 영상: 직접 영상 녹화 및 캡처
- (2) 공개 축구 트래킹 데이터: StatsBomb, Metrica sports sample data 등

3.4 Workflow Diagram



4. 현실적 제약 사항 분석 결과 및 대책

- (1) 실제 경기 트래킹 데이터 확보의 어려움: FM2024를 활용하여 유사한 경기 데이터를 생성, 공개 데이터셋 활용
- (2) 영상 기반 트래킹 정확도 문제: 프레임 간 보간/필터링을 적용하여 정확도 향상, DeepSORT 등 ID 재추적 기술 병행
- (3) 모델 학습 데이터 부족: 데이터 증강 기법 사용, 베이스라인으로 간단한 ML 모델부터 시작 후 점진 확장

5. 개발 일정 및 역할 분담

구분	작업일정																	
	5월			6월			7월			8월			9월					
착수보고서 작성 및 제출																		
FM2024 데이터 취득 및 tracking data전처리																		
모델 Survey 및 선택																		
모델 Tuning 및 최적화																		
중간 보고서 작성																		
시각화 툴 개발																		
결과 분석 및 보수																		
최종 보고서 작성 및 발표 준비																		

5.2 역할분담

이름	역할
최성민	FM2024 영상 캡처 및 CSV, JSON화, 모델 Survey, 모델 학습 및 Tuning, 발표 준비
강동권	FM2024 영상 촬영, 모델 Survey, 시각화 툴 개발, 학습 결과 분석