

GNN 기반 롤 승률 예측 MLOps 프로젝트 기획안

Graph Neural Network를 활용한 리그 오브 레전드 챔피언 선택 단계별
승률 예측 MLOps 서비스 구축 프로젝트

데이터 사이언스 & MLOps 프로젝트

📦 GNN 🏠 ML ⚙️ MLOps

목차

- 1 프로젝트 배경 및 동기
- 2 목표 및 추진 범위
- 3 역할 분담
- 4 단계별 개발 일정
- 5 모델 평가 및 성공 기준
- 6 최종 서비스 형태 및 제공 기능
- 7 기대 효과 및 실무/연구적 가치
- 8 한계 및 리스크



프로젝트 배경 및 동기

매칭 공정성 이슈

최근 매칭메이킹 알고리즘의 공정성에 대한 커뮤니티의 관심이 증가하고 있으며, 챔피언 조합과 승률 간의 객관적 관계에 대한 논의가 활발함

데이터 접근성

라이엇 게임즈의 공개 API를 통해 양질의 대규모 데이터 수집 가능, 일관된 데이터 파이프라인 구축으로 지속적인 모델 업데이트 가능

AI 기술 발전

그래프 신경망(GNN)의 발전으로 챔피언 간 관계, 시너지, 카운터 관계를 효과적으로 모델링할 수 있는 기술적 토대 마련

연구 기회

LoL 도메인에 GNN 적용 선행 연구가 적어 학술적 가치가 높으며, MLOps를 통한 지속적 배포 환경을 구축하는 실무 경험 확보 가능

선행연구 참고

GNN 기반 스포츠/e스포츠 승률 예측 연구가 증가하는 추세로, 2022년 NBA 경기 예측에서 76.9%의 정확도 달성 사례 등 참고 가능

목표 및 추진 범위

🎯 프로젝트 목표

GNN을 포함한 그래프 기반 방법을 활용하여 MLOps를 도입, 챔피언 선택 단계별 "신뢰 가능한 승률 확률"을 제공하는 서비스 구축

✓ 포함 범위

모델 개발 → 프로토타입 → 데스크톱 앱/웹앱 형태의 서비스 제공

✗ 비포함 범위

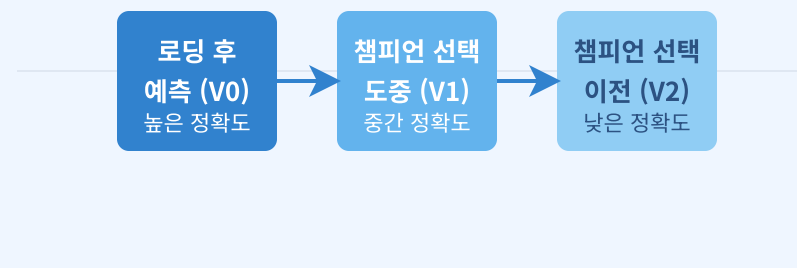
실시간 인게임(경기 중) 예측은 이번 프로젝트 범위에서 제외

📈 기대 효과

투명하고 신뢰할 수 있는 승률 예측으로 게임 경험 향상 및 데이터 기반 의사 결정 지원

3단계 예측 모델

단계별 예측 모델



단계별 접근 전략

V0(로딩 후) 모델을 우선 개발하여 가시적 성과를 확보한 후, V1(선택 도중)과 V2(선택 이전) 모델로 확장하는 점진적 접근 전략 채택

역할 분담

신동빈



팀장/기획/총괄

- ✓ 전체 일정·지표·우선순위 관리, 대외 커뮤니케이션



데이터 사이언티스트

박태현, 조인준

- ✓ EDA, 베이스라인·GNN 연구/개발
- ✓ 전처리 설계, 모델 검증



데이터 엔지니어

박한영, 이준명

- ✓ 데이터 수집·정제·저장 구조 설계
- ✓ 파이프라인/간단 서빙, MLOps 환경 구축 기초

프로젝트로 얻을 수 있는 것

데이터 엔지니어

대규모 수집·정제 파이프라인, 기본 MLOps(실험/버전/배포) 경험

데이터 사이언티스트

실제 서비스 환경에서 동작하는 모델 구축·검증 경험, 대규모 데이터 기반 모델링

단계별 개발 일정

● 공통 ● 데이터 사이언티스트 ● 데이터 엔지니어

2025 2학기 초반부

● 공통

GNN 논문 스터디 및 그래프 구상

일정 세부 계획

● 데이터 엔지니어

소량 데이터 확보·정제 후 데싸 제공

필요 툴·컴퓨팅 환경 조사

데이터/테이블 설계 초안

● 데이터 사이언티스트

롤 데이터 EDA

GNN 심화 학습

전처리 설계

산출물

- EDA 리포트 초안
- 데이터 스키마 초안
- 모델 베이스라인 요구사항

2025 2학기 후반부

● 데이터 엔지니어

테이블/스키마 구축

데이터 적재·전처리 파이프라인 PoC

● 데이터 사이언티스트

베이스라인 모델 설계

초기 그래프 구성 시나리오 수립

산출물

- 베이스라인 지표 1차 결과
- 파이프라인 PoC 동작 데모

2025 겨울방학

● 데이터 엔지니어

서비스 파이프라인 설계

깡통(더미) 모델 연동

● 데이터 사이언티스트

베이스라인 모델 구현/테스트

V0(로딩 후) 프로토타입

산출물

- 내부 데스크톱 데모
- (로딩 후 승률 표출)

2026 이후

● 데이터 엔지니어

서비스 환경 테스트·안정화

● 데이터 사이언티스트

성능 최적화(하이퍼파라미터 튜닝, 칼리브레이션)

V1/V2 단계 확장

산출물

- 베타 서비스
- 단계별 모델 고도화 리포트

모델 평가 및 성공 기준

오프라인 평가 지표

Brier Score

예측 확률과 실제 결과 간의 평균 제곱 오차를 측정하는 점수. 낮을수록 좋은 성능을 나타냄. 목표: 베이스라인 대비 5%+ 개선

Calibration (보정)

예측 확률이 실제 결과 확률과 일치하는지 평가. ECE(Expected Calibration Error)로 측정. 목표: $ECE \leq 0.05$

AUC-ROC

모델의 분류 성능을 평가하는 지표. 1에 가까울수록 좋은 성능을 나타냄. 목표: 0.7 이상

LogLoss

예측 확률의 정확성을 평가하는 손실 함수. 낮을수록 좋은 성능을 나타냄. 보조 지표로 활용

온라인 서비스 지표

응답 지연 시간

P95 기준 100ms 이내 (서버 기준)

가용성/안정성

오류율 0.1% 이하, 타임아웃 0.5% 이하

최종 서비스 형태 및 제공 기능

서버형 모델 아키텍처

GNN 기반 예측 모델을 서버에 배포하여 안정적인 추론 성능 제공. MLflow를 통한 모델 버전 관리와 서빙 자동화로 지속적 업데이트 체계 구축

클라이언트 구성

데스크톱 앱(LCU 연동) 및 웹 기반 싼 클라이언트로 사용자 접근성 최대화. REST API를 통한 예측 요청 및 결과 표출

제공 정보 및 기능

- 1) 승률 예측: V0(선택 후), V1(선택 중), V2(선택 전) 단계별 확률 제공
- 2) 신뢰도 정보: 칼리브레이션 지표로 예측 신뢰성 가시화
- 3) 영향 요인: 승패에 영향을 주는 핵심 요소 요약 제공

확장성 구조

개인화 추천(밴/픽 추천) 기능으로 확장 가능한 모듈형 구조 설계. 인게임 기능은 향후 확장 가능성 열어둠

기대 효과 및 실무/연구적 가치

팀별 학습 경험

데이터 엔지니어링 역량

대규모 데이터 수집/정제 파이프라인 구축, MLOps 기반 CI/CD 환경 구성, 모델 배포 자동화 경험을 통한 실무 역량 강화

데이터 사이언스 심화

GNN 기반 복잡한 관계 모델링, 실제 서비스 환경의 모델 검증 및 칼리브레이션, 대규모 데이터 기반 모델 튜닝 능력 향상

프로젝트 관리 역량

복잡한 ML 프로젝트의 전체 라이프사이클 관리, 애자일 방법론 적용, 팀 간 협업 경험을 통한 협업 프로젝트 관리 능력 배양

연구 및 확장 가능성

학술적 가치

그래프 기반 게임 승률 예측 연구 분야에 기여할 수 있는 새로운 방법론 제시, 논문화 가능성

사업적 확장성

개인화된 챔피언 추천 시스템, 팀 전략 분석 도구, 게임 중계/분석 보조 시스템 등으로 확장 가능

한계 및 리스크

한계 및 리스크

⚠ 모델 성능 미달

기대한 모델 성능에 미치지 못할 경우 서비스 단계 진입이 지연될 수 있음

📈 GNN 학습 곡선

GNN 학습 곡선이 높아 연구 및 튜닝에 예상보다 많은 시간이 소요될 수 있음

🔄 메타 변화

게임 패치로 인한 메타 변화에 따른 모델 재학습 및 성능 저하 가능성

극복 방안

🛡 대안 모델 병행 연구

GNN보다 성능이 우수한 대안 모델 발견 시 전환 가능하도록 병행 연구

🔗 단계적 접근

V0(로딩 후) → V1(선택 도중) → V2(선택 이전) 순으로 복잡도 증가