**TEAM A 프로젝트 제안서**

### 프로젝트명: Spaceship Titanic 승객 위치 예측 모델 구축

## 1.프로젝트 개요

## 서론:Spaceship Titanic의 승객들의 실종 여부를 예측하는 이진분류 모델

## 선택한 이유:흥미로운 주제이면서 다양한 실제 문제에 적용할수있을것이라고 생각했기 때문

## 필요성:Spaceship Titanic은 승객의 다른 차원으로의 실종여부 판단이라는 비현실적인 주제 같아보이지만 프로젝트의 핵심은 주어진 데이터를 기반으로 이진분류 모델을 구축하는 것으로 질병 고위험군 파악, 이탈 고객 예측 등의 다양한 이진 분류 문제에 활용할수있음.

**2. 프로젝트 주제 및 데이터셋**

**문제 정의**

주어진 데이터로부터Spaceship Titanic의 실종된 승객들 중 누가 다른 차원으로 이동되었는지 이진분류 모델을 이용해 예측

**데이터 셋**

1.Sample\_submission데이터 크기: 59KB / 9277 건

2.train데이터 크기: 787KB / 9280 건

3.test데이터 크기: 364KB / 9277 건

**주요 특성**

**입력 변수**: [승객이 출발한 행성, 냉동수면여부, 객실번호, 승객이 도착할 행성, 나이, VIP여부, 편의시설사용요금, 이름]

**타겟 변수**: [승객의 실종 여부]

**3.과정**

1. 데이터 탐색: 데이터 구조와 변수 파악, 시각화로 분포 및 상관관계 분석.

2. 데이터 전처리: 결측치 및 이상치 처리, 스케일링 및 인코딩.

3. 특징 엔지니어링: 중요한 변수 선택 및 새로운 변수 생성.

4. 모델 학습: 적합한 모델 선택, 학습 및 하이퍼파라미터 튜닝.

5. 모델 평가: 교차 검증, 평가지표로 성능 분석.

6. 최종 테스트 및 보고: 최적 모델로 예측, 결과 시각화 및 보고서 작성.

**팀원 개별역할**

2015108전현태:팀 리더, 데이터 이해, 결측값 처리, 데이터 분석 및 전처리

2226046우지은:PTT제작, 모델링 및 하이퍼파라미터 튜닝

2126076함서현:발표, 모델 성능 분석, 시각적 자료 작성