

1. 탐색적 데이터 분석 (Exploratory Data Analysis, EDA)

목표:

데이터를 시각화하고 통계적으로 분석하여 데이터의 구조, 특성, 패턴을 이해하고, 분석할 만한 인사이트를 도출하는 것이 목표입니다.

내용:

- 데이터 분포:
채용 공고의 직무, 지역, 요구 기술, 경력 수준 등 주요 변수들에 대한 분포를 확인하여 전체 데이터의 구조를 파악합니다.
- 통계적 요약:
평균, 중앙값, 표준편차 등의 기초 통계값을 구해 채용 공고의 주요 변수들(연봉, 경력 연수 등)의 특징을 이해합니다.
- 시각화:
 - 히스토그램, 막대그래프, 상자 수염 그림(box plot) 등을 통해 변수의 분포를 시각화합니다.
 - 연봉 분포, 기술 스택별 채용 현황, 직무별 공고 수 등을 시각적으로 분석하여 주요 트렌드나 패턴을 파악합니다.
- 이상치 탐지:
데이터 내 이상치나 결측값을 탐지하여 이후 분석에 방해가 되지 않도록 처리합니다.

기법 활용 예시:

- 직무별 채용 공고의 개수와 평균 연봉 간의 관계를 시각화하여 어떤 직무에 높은 연봉이 제공되는지 확인.
- 특정 기술(예: "Python")이 요구되는 채용 공고 수와 다른 기술들과의 연관성 파악.

2. 상관관계 분석 (Correlation Analysis)

목표:

변수들 간의 관계를 파악하여, 어떤 변수들이 서로 밀접하게 연관되어 있는지 확인합니다. 이를 통해 직무, 연봉, 경력 수준, 요구 기술 등의 관계를 분석할 수 있습니다.

내용:

- 상관 행렬(Correlation Matrix):

변수 간의 상관 계수를 계산하여 어떤 변수들이 서로 상관관계가 높은지 시각적으로 나타냅니다. 예를 들어, 연봉과 경력 연수, 특정 기술의 요구와 직무 간의 관계를 파악할 수 있습니다.

- 산점도 및 회귀 분석:

주요 변수 간의 관계를 산점도를 이용해 시각화하고, 이를 기반으로 선형 회귀 분석을 통해 변수 간의 예측 관계를 모델링합니다.

- Pearson, Spearman 상관 계수:

Pearson 상관 계수를 사용하여 두 변수 간의 선형 상관 관계를 확인하고, Spearman 상관 계수를 통해 비선형 관계를 분석할 수 있습니다.

기법 활용 예시:

- "경력 연수"와 "연봉" 간의 상관 관계를 분석하여 경력이 많을수록 연봉이 증가하는 패턴을 확인.
- "기술 스택"과 "직무" 간의 관계를 분석하여 특정 기술이 요구되는 직무의 비율을 파악.

3. 머신러닝 모델 개발 (Machine Learning Model Development)

목표:

채용 공고 데이터를 기반으로 예측 모델을 개발하여, 구직자에게 맞춤형 채용 공고를 추천하거나, 채용 공고의 주요 특징을 예측합니다.

내용:

- 분류 모델 개발:

채용 공고를 분류하는 모델을 개발합니다. 예를 들어, 채용 공고를 "개발자", "디자이너", "마케팅" 등으로 분류할 수 있습니다.

- 알고리즘 예시:

- 로지스틱 회귀(Logistic Regression)
- 결정 트리(Decision Tree)
- 랜덤 포레스트(Random Forest)
- 서포트 벡터 머신(SVM)

- 회귀 모델 개발:

예측 변수(예: 연봉, 근무 시간 등)를 예측하는 회귀 모델을 구축할 수 있습니다. 예를 들어, 특정

직무의 연봉을 예측하는 모델을 개발할 수 있습니다.

- 알고리즘 예시:
 - 선형 회귀(Linear Regression)
 - 다항 회귀(Polynomial Regression)
- 추천 시스템:

구직자의 이력서와 채용 공고의 텍스트 데이터를 기반으로 맞춤형 채용 공고를 추천하는 모델을 개발합니다.

 - 알고리즘 예시:
 - 협업 필터링(Collaborative Filtering)
 - 콘텐츠 기반 필터링(Content-based Filtering)
 - 하이브리드 추천 시스템

기법 활용 예시:

- 분류 모델:

직무를 분류하는 모델을 사용하여 채용 공고가 어떤 분야에 속하는지 예측합니다. 예: "개발자" 직무에 해당하는 공고를 자동으로 분류.
- 회귀 모델:

직무별 평균 연봉을 예측하는 모델을 개발하여, 채용 공고에서 예상 연봉을 추정합니다.
- 추천 시스템:

구직자의 이력서를 기반으로, 가장 적합한 채용 공고를 추천합니다.

결론

이 세 가지 분석 기법을 통해 채용 정보 분석 플랫폼은 다음과 같은 목적을 달성할 수 있습니다:

1. ****탐색적 데이터 분석(EDA)****를 통해 데이터의 전반적인 특성과 패턴을 이해하고, 중요한 인사이트를 도출.
2. 상관관계 분석을 통해 변수 간의 관계를 분석하고, 데이터에서 중요한 관계를 파악하여 예측 모델에 활용.
3. 머신러닝 모델 개발을 통해 실시간으로 구직자에게 맞춤형 채용 공고를 추천하거나, 채용 공고의 주요 특성에 대한 예측을 수행.

이러한 분석 기법들은 플랫폼의 데이터 분석 정확도와 사용자 맞춤형 서비스를 개선하는 데 중요한 역할을 할 것입니다.