

# Artificial Intelligence Project for Data Mining



Ali Eskandarian / Deadline:

## Project Description

در این پروژه شما باید رزومه های فارسی ارسال شده برای استخدام در حوزه ی هوش مصنوعی را بررسی نموده و درصد تطبیق رزومه های ارسال شده را محاسبه کنید. برای محاسبه ی درصد تطبیق رزومه ها باید به اطلاعات نوشته شده در آگهی استخدامی آن مراجعه کنید. برای مثال: درصد تطابق رزومه ی A با آگهی کارآموزی هوش مصنوعی ۶۵٪ می باشد.

برای انجام این پروژه میتوانید از آگهی کارآموزان هوش مصنوعی شرکت تاو سیستم استفاده کنید. پارامتر های مهم موجود در آگهی را در نظر بگیرید (به عنوان مرجع برای سنجش تطابق) و با توجه به اطلاعات نوشته شده در رزومه ها، درصد تطبیق هر رزومه با آگهی را مشخص کنید. در خروجی باید بتوانیم مشخصات رزومه های که بیشترین تطابق را دارند را مشاهده کنیم.

منظور از تطابق رزومه ها با آگهی استخدامی چیست؟ کارفرما در آگهی استخدامی خود شرایطی را برای استخدام افراد عنوان میکند (مثلا افراد بالای 20 سال، ساکن شهر شیراز، مهارت متوسط در برنامه نویسی پایتون و ...). حال شما باید درصد تطبیق رزومه ها را با شرایط گفته شده در آگهی محاسبه کنید. و در نظر داشته باشید که متن آگهی استخدام باید داینامیک باشد.

- اصول clean code را در کد های خود رعایت کنید.
  - در تاریخ تحویل پروژه لینک گیت هاب خود را که کدها و دیتا هایتان در آن قرار دارند را برای ما ارسال کنید.
- xx توجه:** نباید از مدل های license دار تجاری استفاده کنید.

## API Description ❖

- داده های ورودی که به API شما داده میشود:
  1. متن آگهی استخدام - string
  2. یک فایل متنی رزومه ی فارسی
- خروجی مورد انتظار که باید توسط API بازگردانده شود:

برای هر رزومه یک فایل json نمایش داده شود که شامل موارد زیر باشد:

  1. درصد تطابق رزومه با آگهی،
  2. نام و نام خانوادگی ارسال کننده ی رزومه،
  3. ویژگی هایی از رزومه ی فرد که برای سنجش تطابق استفاده کردید.

به عنوان مثال خروجی باید به شکل زیر باشد \_ (نام گذاری ها و انتخاب کلیدها به عهده ی شماست) :

```
{  
  "Similarity": '30',  
  "Name": 'negar',  
  "Lastname": 'negahdaripour',  
  "Age": '30',  
  "Location": 'shiraz',  
  "Python_skill": 'intermediate',  
  "Git": 'beginner',  
  Other ...  
}
```

شما باید تمام مراحل انجام این پروژه را داکيومنت کنید (۲-۳ صفحه کفایت میکند) و توضیحات شما باید شامل سه بخش باشد:

### ❖ Data Preparation

در این بخش باید در مورد نوع داده هایی که استفاده کرده اید، روشی که برای جمع آوری داده ها انتخاب کردید، مرحله ای که برای clean کردن داده ها طی کردید، نحوه ی ساختن دیتاست خود و روش هایی preprocessing خود و چالش هایی که در این بخش داشتید توضیح دهید.

### ❖ Data Mining Algorithm

در این بخش شما باید درمورد الگوریتمی که انتخاب کردید توضیح دهید. (مواردی مثل ویژگی های الگوریتم و زمینه ی کاربرد آن، علت انتخاب این الگوریتم، روش کارکرد آن و چالش هایی که شاید با آنها مواجه شده باشید).

### ❖ Your results

نتایجی که به دست آوردید را شرح دهید.

**“The only way to do great work is to love what you do.” Steve Jobs**