Eightfold.ai Machine Learning Engineering Take Home Project

Thank you for taking the time to take on the eightfold.ai engineering challenge! We use this project as a way to understand your coding abilities as well as how you would work within a team here at Eightfold.

We have designed the project to emulate our development process here at Eightfold. You will receive a design and a specification (spec) document. You have full freedom to interpret the spec and design and build the feature out end to end. You can use any external libraries that you are most comfortable with for training the models. When you submit your code, we are not just looking for code that works but for code that is production ready and is able to be merged into our codebase.

You will be building a seniority classifier. Given a profile's education and experience history, classify their seniority into the following buckets

1. Intern
2. Entry
3. Mid Level
4. Senior
5. Manager
6. Director
7. Vice President
8. CXO

You will be provided with training data and testing data via a separate shared google drive. Each row in the training data contains experience, education history and seniority level. The seniority level is the label you would predict.

Sample Row:


```
{
        "seniority_level": "Vice President",
        "education": [{
                "school": "William E. Simon School of Business Administration Rochester",
                "degree": "MBA",
                "major": "Finance & Marketing",
                "time": ["notKnown", "1996-01-01"],
```

```
                "description": "William E. Simon School of Business Administration (Top 20 B-School / Top 10 Finance)\nUniversity of Rochester, Rochester, New York\nMBA, 1996, Finance & Marketing, Merit-Based Scholarship"
        }, {
                "school": "Trinity College Hartford",
                "degree": "Bachelor of Arts",
                "major": "Economics",
                "time": ["notKnown", "1991-01-01"],
                "description": "Trinity College, Hartford, CT\nBachelor of Arts, 1991, Economics, Merit-Based Scholarship - Trinity Club"
        }
        …
        ],
        "experience": [{
                "description": "Business Partnering\nOperations Reviews\nAnnual Plan\nDash Boards\nP&L Forecasting\nReporting\nBusiness Cases\nSystems\nMonthly Close\nDirector\nCisco Systems\n(Public Company; Computer Networking)",
                "title": "Vice President of Finance",
                "skills": ["cisco", "business operations", "operations", "networking", "dash", "reporting", "dashboards", "reviews", "monthly close"],
                "work": "Self Employed",
                "time": ["2013-12-01", "Present"]
        }, {
                "description": "Member of Executive Management Team who led all finance activities for recruiting firm that places candidates across the US and specializes in Sales, Engineering and Finance.\n\u25cf Business partner on all finance projects.\n\u25cf Key participant in all business development activities to win new clients in SF and Silicon Valley.\n\u25cf Owned the business relationship with Huawei, a Chinese Networking company with revenue of $60B.  Champion placed 14 executives with Huawei in primarily Sales Engineer and Solution Architect positions.",
                "title": "Vice President of Finance",
                "skills": ["finance", "management", "sales", "networking", "silicon", "revenue", "clients", "business development", "sales engineering", "engineering", "executive management", "projects", "engineer"],
                "work": "Champion Recruiting",
                "time": ["2013-07-01", "2013-12-01"],
        },
        …
        ]
}
```

Evaluation Criteria:

For this project you will be evaluated on the accuracy and the quality of your code. The submission of the project should be something you feel really proud of. Please feel free to use any open source library/framework, and please submit your project as a GitHub or Bitbucket repo or raw files with instructions on how to run your project. Please also include assumptions and tradeoffs you made while building out the component. We have designed our project to be completed under 3 hours, however, feel free to take up to 3 days to complete the project.

Data Processing

You are provided with training and testing datasets. Build a component in your submission that handles the data processing and sanitization.

Feature Engineering

For this project, build 3-4 features (to be discussed during the briefing session) that might assist the model. Ensure the framework is general purpose so adding new features is easy and intuitive.
For the feature engineering, we look for the following criteria as we review your code

1.) Does your code work? Are all features in working condition?
2.) Runtime Efficiency.
3.) Code Readability.
4.) Error Handling.

Bonus: If you have ideas for additional features feel free to implement them.

Model Training

Now that you have features, feel free to implement a seniority classifier. For the model be specific about the tradeoffs for the choice of the model. Also, include evaluation criteria for the performance of your model.