

## XGBoost — Performance on 20% Unseen Text Data

Accuracy: 98.33%

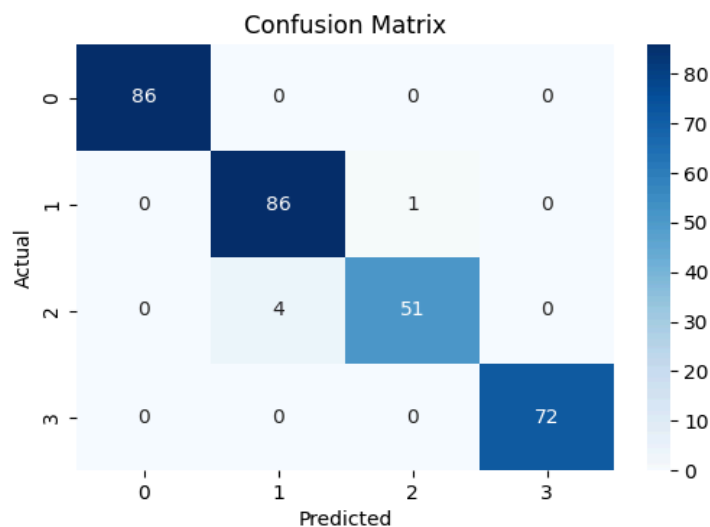
Precision: 98.36%

Recall: 98.33%

F1 Score: 98.32%

Confusion Matrix (Class Mapping):

- 0 → Backend
- 1 → Frontend
- 2 → Fullstack
- 3 → QA



### Major Failure Modes

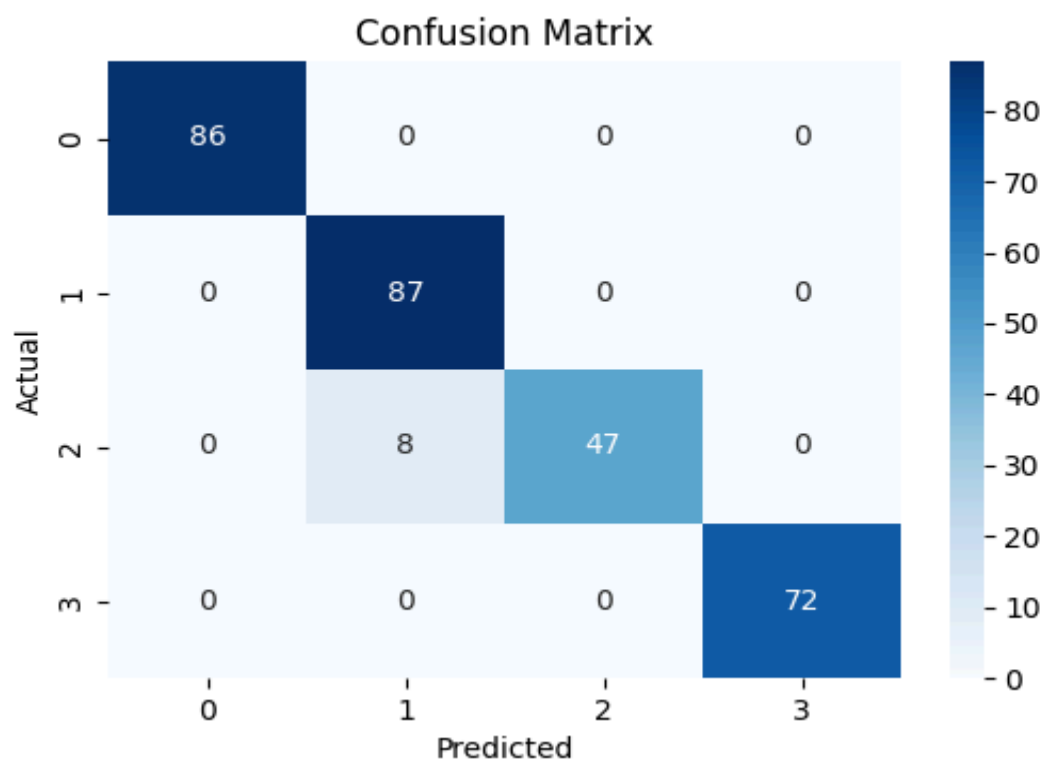
1. Misclassifies **Frontend vs Fullstack** due to overlapping skillsets
2. Model might be overfitting as the data is little low for some classes

## Random Forest

- Accuracy: 0.9733

- **Precision:** 0.9756
- **Recall:** 0.9733
- **F1 Score:** 0.9729

**Confusion Matrix (labels: 0=Backend, 1=Frontend, 2=Fullstack, 3=QA):**



## Major Failure Modes

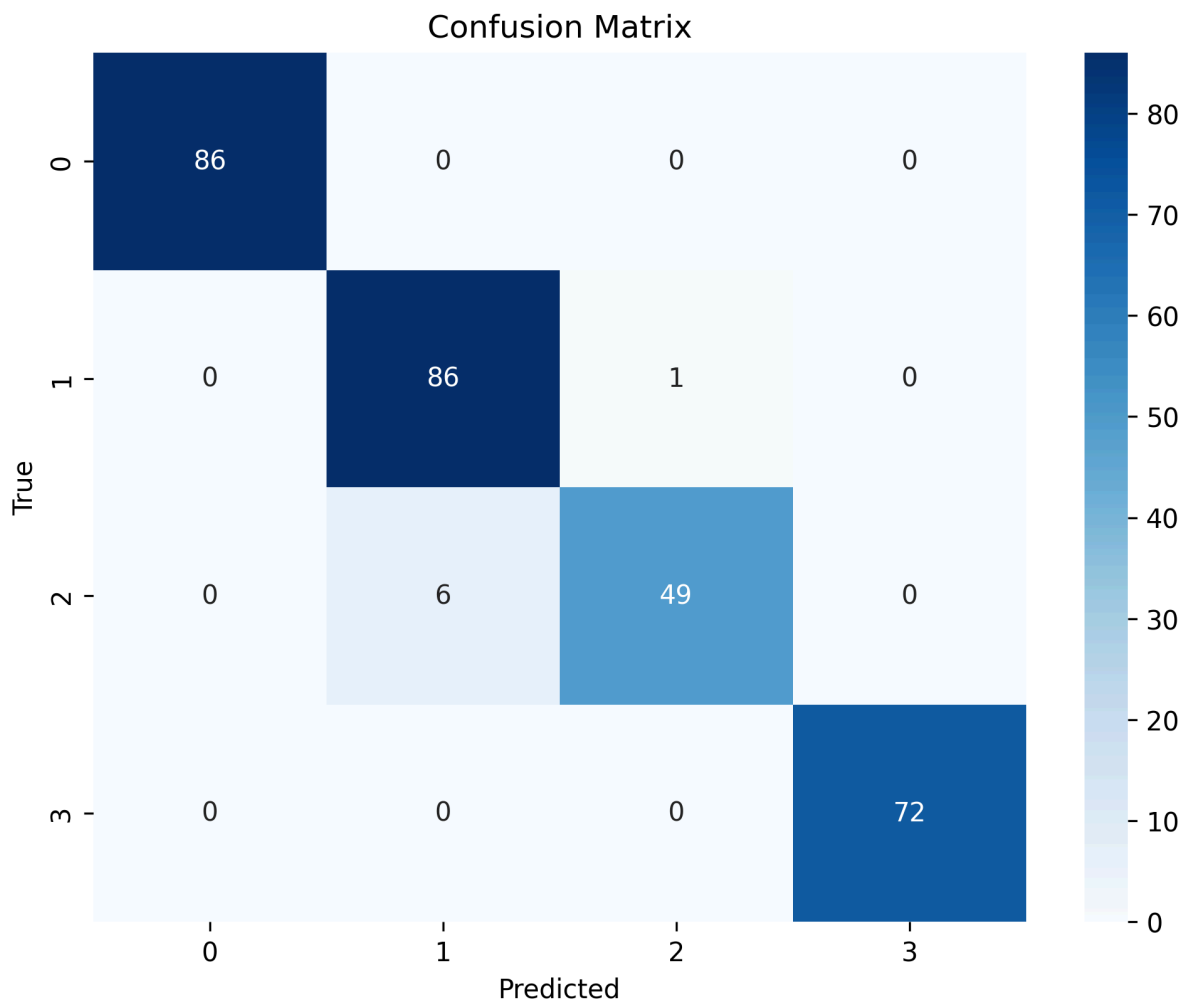
1. Misclassifies **Frontend vs Fullstack** due to overlapping skillsets

## BiLSTM — Performance on 20% Unseen Text Data

- **Accuracy:** 0.9767

- **Precision:** 0.9774
- **Recall:** 0.9767
- **F1 Score:** 0.9764

**Confusion Matrix (labels: 0=Backend, 1=Frontend, 2=Fullstack, 3=QA):**



This is also showing similar behaviour, confusing frontend with the full stack.

### LESSONS:

1. **TF-IDF is a great preprocessing technique and helped us reach very high accuracy, but it is still making mistake on capturing context**
2. **As the data is low, the model is highly likely to overfit**
3. **The fact that BILSTMS do not improve accuracy, shows that the context was not important here.**

