Fine Tuning NER to Extract Financial Entities

We have a general purpose NER model which is not effective in extracting the financial entities that we want to extract from the document. For this purpose we need to fine tune our general purpose NER model on the domain specific dataset for it to accurately extract the entities that we want to. The steps to follow are as follow -

1. Data Preparation - We need data to retrain our NER model on our domain specific data for it to correctly extract the financial entities. We need to collect data from chat logs, emails, trading notes etc and the data should be cleaned for sensitive data and also anonymize data which can't be shared with others. If data cannot be gathered from sources, we need to manually create the data.

2. Data Annotation - We need to label our data that we have collected with correct entity tags. For example - "Bank ABC" - Counterparty, "200 mio" - Notional etc. This will become the dataset on which we will fine tune our general purpose NER model. We also need to perform BIO Tagging as the NLP model requires token-level labels. This ensures the model can recognize multi-word entities like "BANK ABC" or "2Y EVG".

3. Data Splitting - This is a very common stage for any model training, retraining or fine tuning. We split the data into a train, test & validation set. This is important as if not done correctly, the model may memorize things instead of learning. The training set helps the model learn, the validation set helps us tune the hyperparameters & test set helps to check the performance of the model. Split into train (80%), validation (10%), test (10%).

4. Model Selection - As we are not training the model from scratch, we'll choose a pre-trained model which we'll fine tune, like"dslim/bert-base-NER" or domain specific model like FinBERT. Then, add a classification head for token classification (NER). Pre-trained models already understand English structure, so you only need to teach them financial entities.

5. Fine Tune - Train the model on your dataset with hyperparameters like: Learning rate, Batch size, Epochs. The model gradually learns to tag tokens with the correct entity label. Fine-tuning aligns a general-purpose model with your specific financial use case.

6. Testing & Evaluation - Feed unseen text and see if the model extracts entities correctly. We can evaluate using metrics such as precision, F1-Score, Recall etc. Models sometimes miss obvious patterns (e.g., ISINs always follow a specific regex).

We can also use LoRA/QLoRA, if we choose some mid size or large model where retraining would not be computing resource friendly.