Building a Personally Identifiable Information Recognizer in a Privacy Preserved Manner using Automated Annotation and Federated Learning

Submitted by: Rajitha Hathurusinghe
Supervisors: Prof. Miodrag Bolić and Dr. Isar Nejadgholi

Faculté de génie | Faculty of Engineering
uOttawa.ca

Outline

- Problem statement
- Summary of Literature
- Methodology
- Conclusions
- Summary of Contributions

Problem setup

Training deep learning models for recognizing Personally Identifiable Information(PII) in unstructured text.

- Restrictions due to the privacy and sensitive nature of data for:
  - -Collection
  - -Annotation

Applications

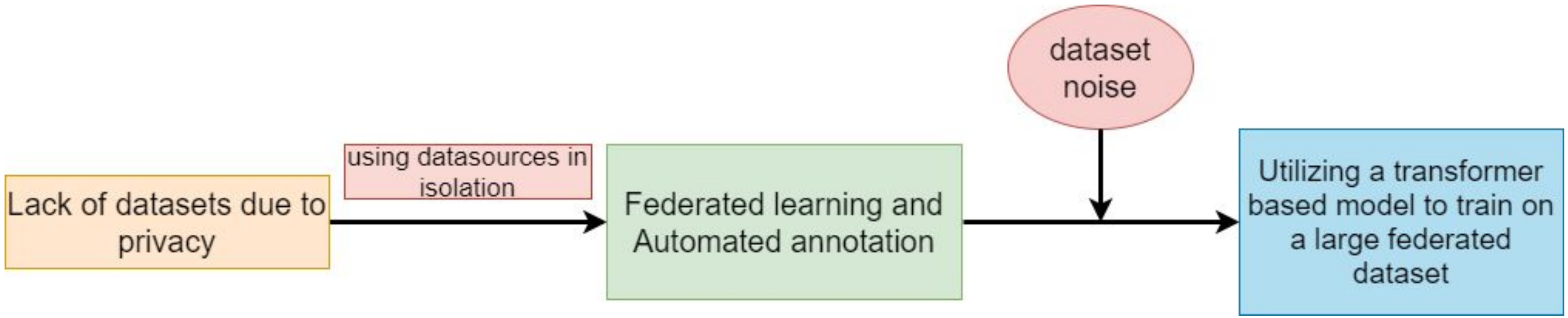
- De-identification of documents
  - Health records,
  - Access to Information and Privacy (ATIP) Online Requests

- Extraction of entities for indexing

Existing solutions

- Differential privacy
  - does not address consent issues with private data
  - effective on privacy of queries on data
- Manually de-identified datasets
  - risk of exposure and requirement of consent
  - famous cases of re-identification
  - cost

Implications and solution



Methodology

- Creation of WikiPII dataset
  - Automated annotation of Wikipedia biography pages
  - Evaluating the dataset comparing to manual annotations

- Fine-tuning BERT-base model in central and federated settings
  - Training in central and remote execution scenarios
  - Training with different volumes of dataset by increasing the number of remote workers
  - Investigating the impact of federated learning on BERT-base model with a popular dataset, CoNLL2003

Contribution 1: WikiPII dataset

- Extracting private entities from the info box
- Fuzzy string matching of extracted private entities on text for annotation

Results: WikiPII dataset

dataset	Entries	sentences	BD	PR	SP	CH	ED
training	20039	77703	16883	6326	25163	10824	24365
validation	2744	12267	2512	1509	3844	1846	3831
test	307	2051	303	331	609	604	534
test (manual)	91	320	76	50	80	62	92

Details of the dataset

Results: Comparison and training performance

	type	partial	strict	exact
precision	0.46	0.39	0.31	0.32
recall	0.65	0.57	0.45	0.46
f1	0.54	0.47	0.37	0.38

Comparison of annotations with manual annotations

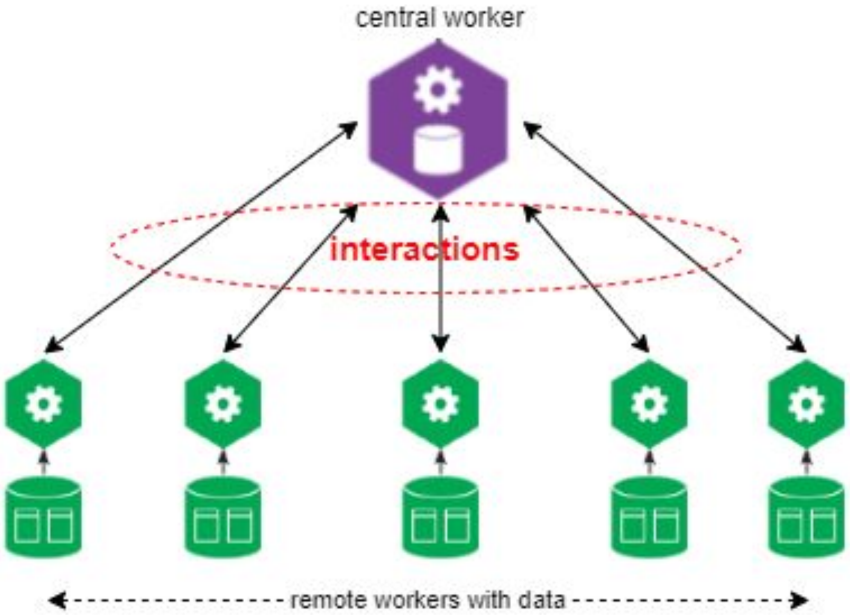
	type	partial	strict	exact
precision	0.79	0.68	0.55	0.56
recall	0.80	0.68	0.56	0.56
f1	0.80	0.68	0.56	0.55

Performance of the trained model on manual test set

uOttawa

Contribution 2: Federated Learning of a NER task with BERT-base NER model

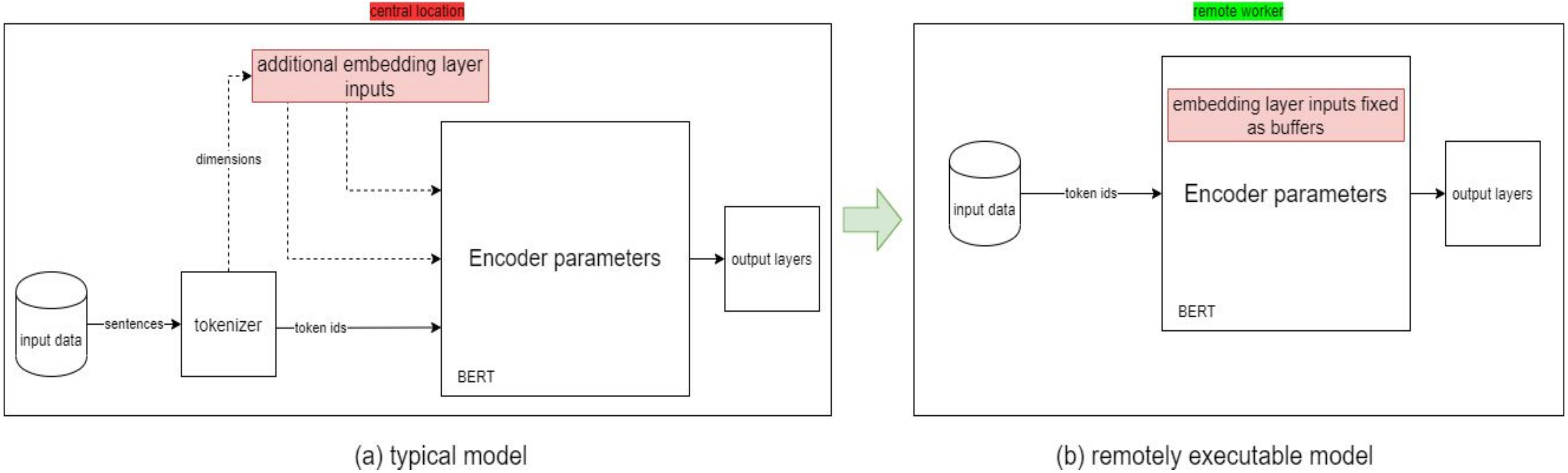
- Utilizing ***PySyft*** framework to create federated dataset and training setup
  - Federated-central
  - Federated-remote



Federated learning setup

Contribution 2: Federated Learning of a NER task with BERT-base

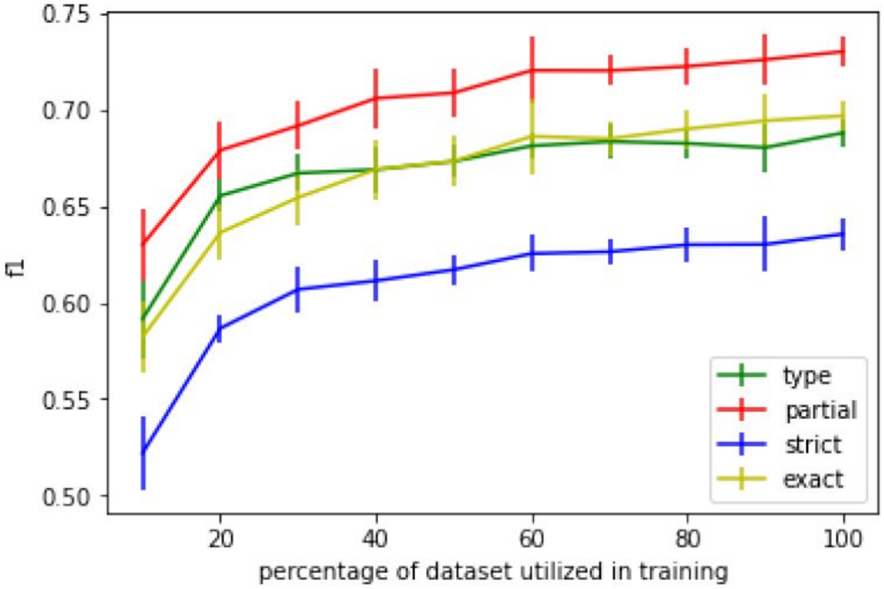
Enabling the transfer of non-parameter tensor buffers needed by the model



Extending BERT-base model for remote execution with PySyft

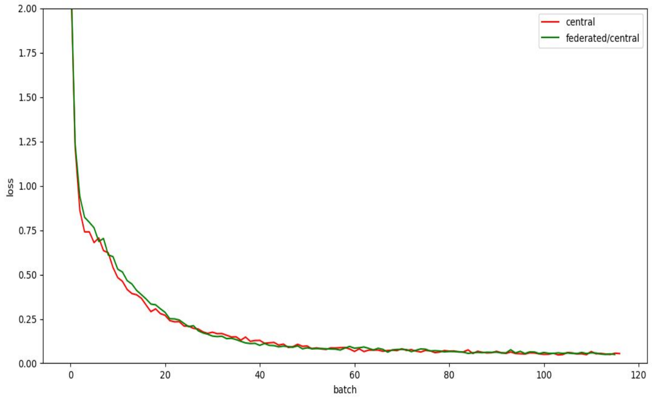
Results: Federated learning

Impact on final performance
when increasing the
data volume
by adding federated datasets.

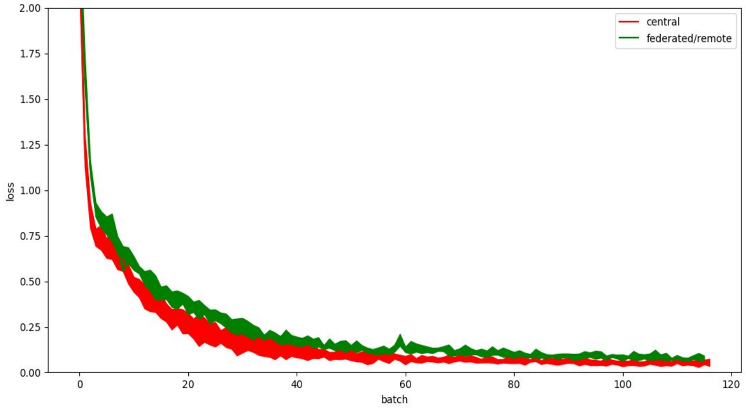


Results: Federated learning

Training loss compared to that of non-federated training



Centrally operated model



Remotely operated model

Results: Federated learning

dataset	training setting	no. of workers	F1 score
CoNLL2003	central	N/A	0.90 $\pm$0.005
	federated/remote	2	0.85 $\pm$0.003
	federated/central	2	0.90 $\pm$0.008
WikiPII	central	N/A	0.70 $\pm$0.006
	federated/remote	2	0.56 $\pm$0.02
	federated/central	2	0.70 $\pm$0.01

Training performance of the model for all the training settings

Conclusions

- Even with a simple rule set a substantially big dataset can be created **inexpensively** to train a NER model and gain a good accuracy in a PII recognizing task overcoming the annotation **noise**.

- Federated learning can be used to increase the volume of training data then gain accuracy of the transformer based models on NER while preserving the privacy of the data sources.

- Remote execution by weight transfer of the model has an impact on the final accuracy of the model.

Summary of contributions

1. WikiPII dataset with PII data

2. Extension of BERT-base model for remote execution and required developments

3. Impact of weight transfer on BERT-base model upon federated learning: inability to minimize training error leading to a lower accuracy

4. Proof of the concept of using automated annotation and federated learning for training a PII recognizer

Thank You