

IndicFed: A Federated Approach for Sentiment Analysis in Indic Languages

Jash Mehta*, Deep Gandhi*, Naitik Rathod, Sudhir Bagul[†]

Dwarkadas J. Sanghvi College of Engineering
Mumbai, India

{jashmehta3300, thisisdeepgandhi, naitikrathod18}@gmail.com

[†]sudhir.bagul@djsce.ac.in

A Appendix

A.1 Basic Task Formulation

Sentiment analysis is a type of text classification task where the input is a sequence of words, $x = [w_1, w_2, \dots]$, and the output is a label y in a fixed set of labels \mathcal{L} . Here, $\mathcal{L} \in \{pos, neg, neu\}$

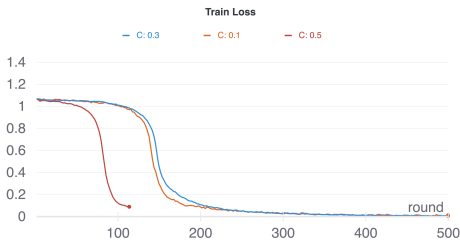
A.2 Server-Based Models

For traditional server-side transformer models, we use the `simpletransformers`¹ library. We use the default configuration options. We train all the transformer models for 25 epochs with a learning rate of $4e-5$ and no weight decay or momentum. Sentences are processed in batches of 8. The input sentences were tokenized with the maximum sequence length restricted to 128.

A.3 Plots for Federated models

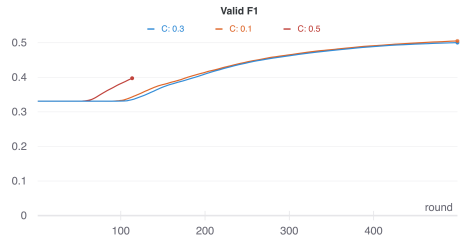
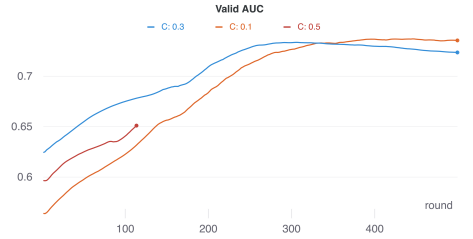
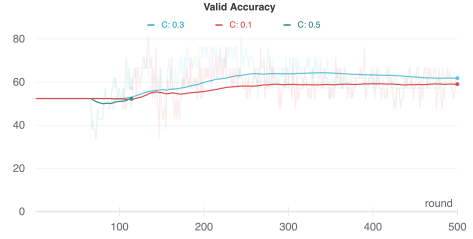
In this section, we show some of the plots for our experiments for each language. These plots clearly show the different models that we train as part of our training setting described in the paper. We use Weights and Biases (Biewald, 2020) to track the performance. In all the plots, c represents the fraction of randomly selected clients for every round.

A.3.1 Telugu (IID)



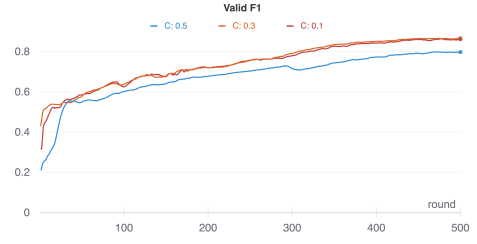
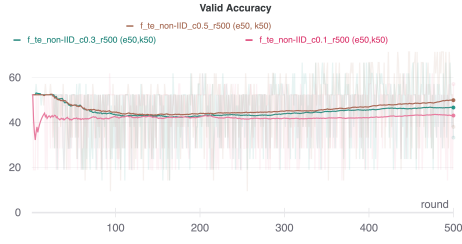
* indicates equal contribution

¹<https://simpletransformers.ai/>

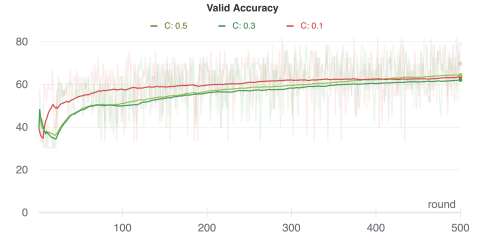
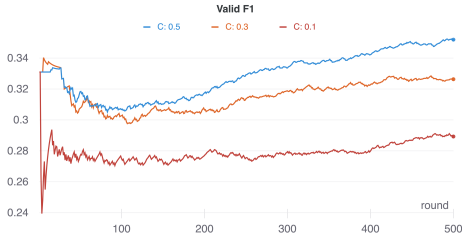
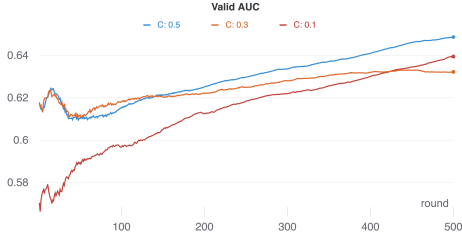


A.3.2 Telugu (non-IID)

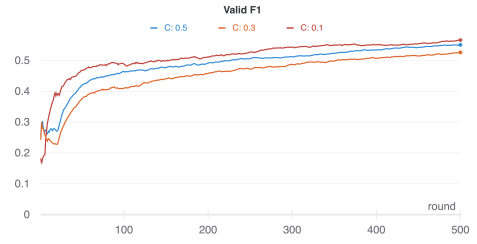
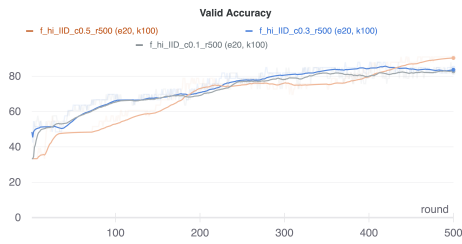
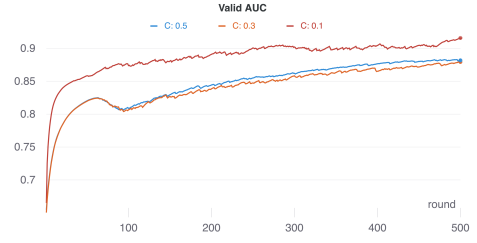




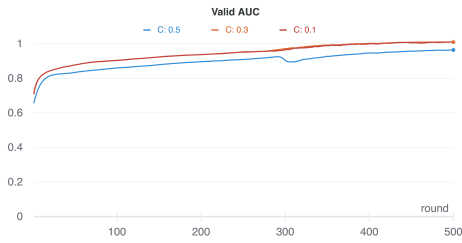
A.3.4 Hindi (non-IID)

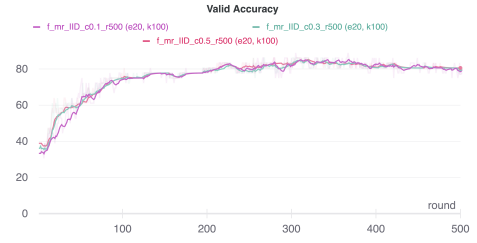
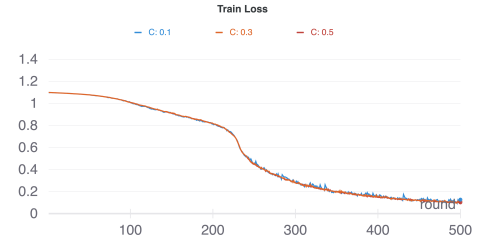
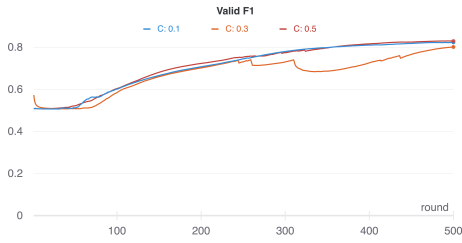
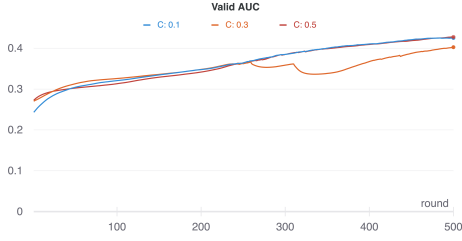
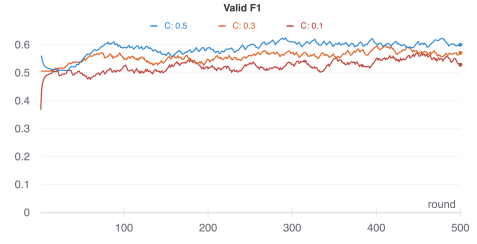
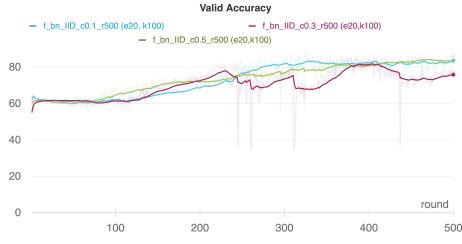


A.3.3 Hindi (IID)

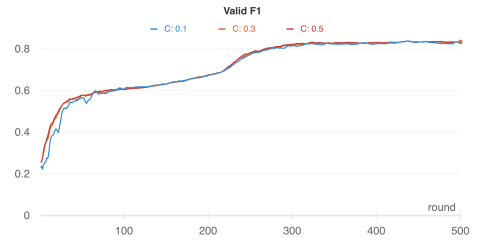
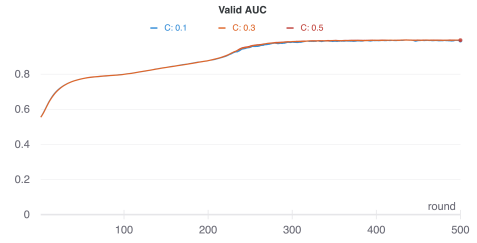
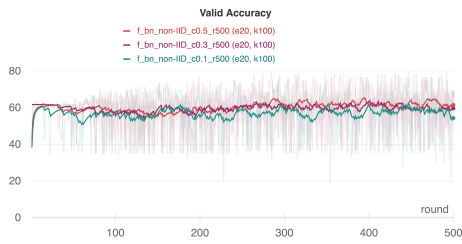


A.3.5 Bengali (IID)

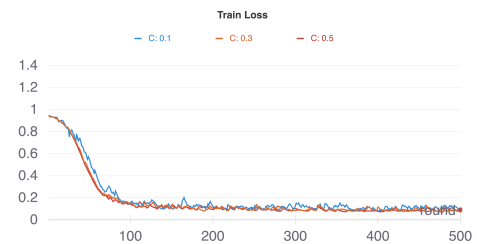
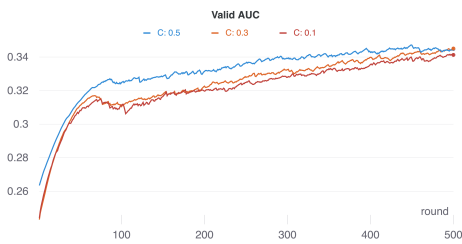


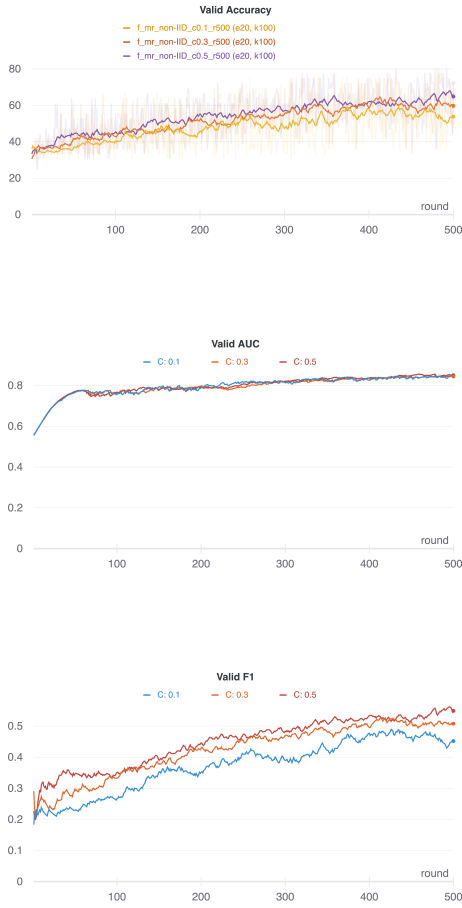


A.3.6 Bengali (non-IID)

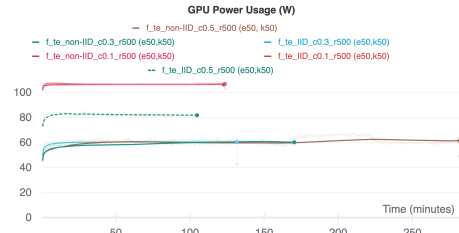


A.3.8 Marathi (non-IID)

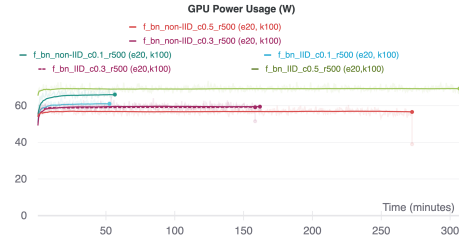




A.4.3 Telugu



A.4.4 Bengali



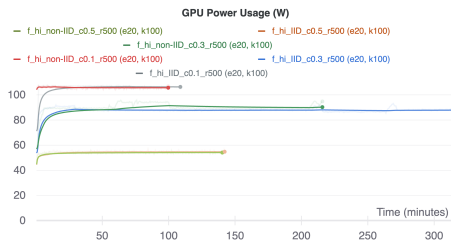
References

Lukas Biewald. 2020. [Experiment tracking with weights and biases](https://wandb.com). Software available from wandb.com.

A.4 Time vs GPU Usage

This section provides detailed graphs for GPU usage in Watts for every variation of experiments run.

A.4.1 Hindi



A.4.2 Marathi

