

Federated Learning by Stacking

M1 Supervised Project
Part 1: Bibliographic Report

Submitted in partial fulfillment
of the requirements for the degree of
MSc Natural Language Processing

by

Anna MOSOLOVA
Elisa LUBRINI
Pierrick FOULQUIER

Supervisor:

Christophe Cerisara

Reviewer:

...



Academic year 2020-2021

Contents

1	Introduction	1
2	Theoretical Background	2
2.1	Ensemble learning	2
2.1.1	Meta-learning and stacked generalisation	2
2.1.2	Stacking in NLP	3
2.2	Federated Learning	3
2.2.1	One-shot Federated learning	3
2.2.2	(One-shot?) federated learning in NLP	3
3	Current problems	4
4	Prior art	5
4.1	Benchmarks	5
4.1.1	SentEval	5
4.1.2	SuperGLUE	5
4.1.3	RoBERTa	5
	Bibliography	6

Chapter 1

Introduction

This report represents a bibliographic analysis of state-of-the-art techniques for stacking ensemble methods and federated learning, particularly one-shot federated learning. The purpose of this report is recording our search for a meeting point for stacking and federated learning methods to be applied in the resolution of a common deep learning problem.

[...]

Chapter 2

Theoretical Background

2.1 Ensemble learning

Ensemble learning is a machine learning technique that consists in combining multiple models, henceforth referred to as base-learners, which have been previously trained to solve the same problem. Strategically combining their results aims at improving predictions and obtain an overall better performance compared to that of each of the single models.

Ensemble learning is considered to be based on the idea of "wisdom of the crowd"[1]. This concept can be easily illustrated with an experiment conducted by Francis Galton in the 19th century. The English philosopher and statistician, also known for conceiving the concepts of standard deviation and correlation, is said to have organised a contest during a livestock fair, in which he asked various people to guess the weight of an ox. Although none of the participants guessed the exact weight, the average of the predictions was quite close to the actual value. This experiment proved the power of combining different guesses to achieve a more accurate prediction, an idea subsequently implemented in machine learning through ensemble modelling.

2.1.1 Meta-learning and stacked generalisation

One of the best performing techniques in ensemble modelling consists in developing an architecture that assembles more than one learning stage, in which base-learners' output is reprocessed by meta-learners in an effort to discern which models are stronger and which make inaccurate predictions.

[...]

Stacking is likely to be the most prominent meta-learning technique[2]

[...]

2.1.2 Stacking in NLP

2.2 Federated Learning

2.2.1 One-shot Federated learning

2.2.2 (¿One-shot?) federated learning in NLP

Chapter 3

Current problems

Chapter 4

Prior art

4.1 Benchmarks

4.1.1 SentEval

4.1.2 SuperGLUE

4.1.3 RoBERTa

Bibliography

- [1] O. Sagi and L. Rokach. Ensemble learning: A survey. *WIREs Data Mining and Knowledge Discovery*, 8(4):e1249, 2018.
- [2] D. H. Wolpert. Stacked generalization. *Neural Networks*, 5(2):241 – 259, 1992.