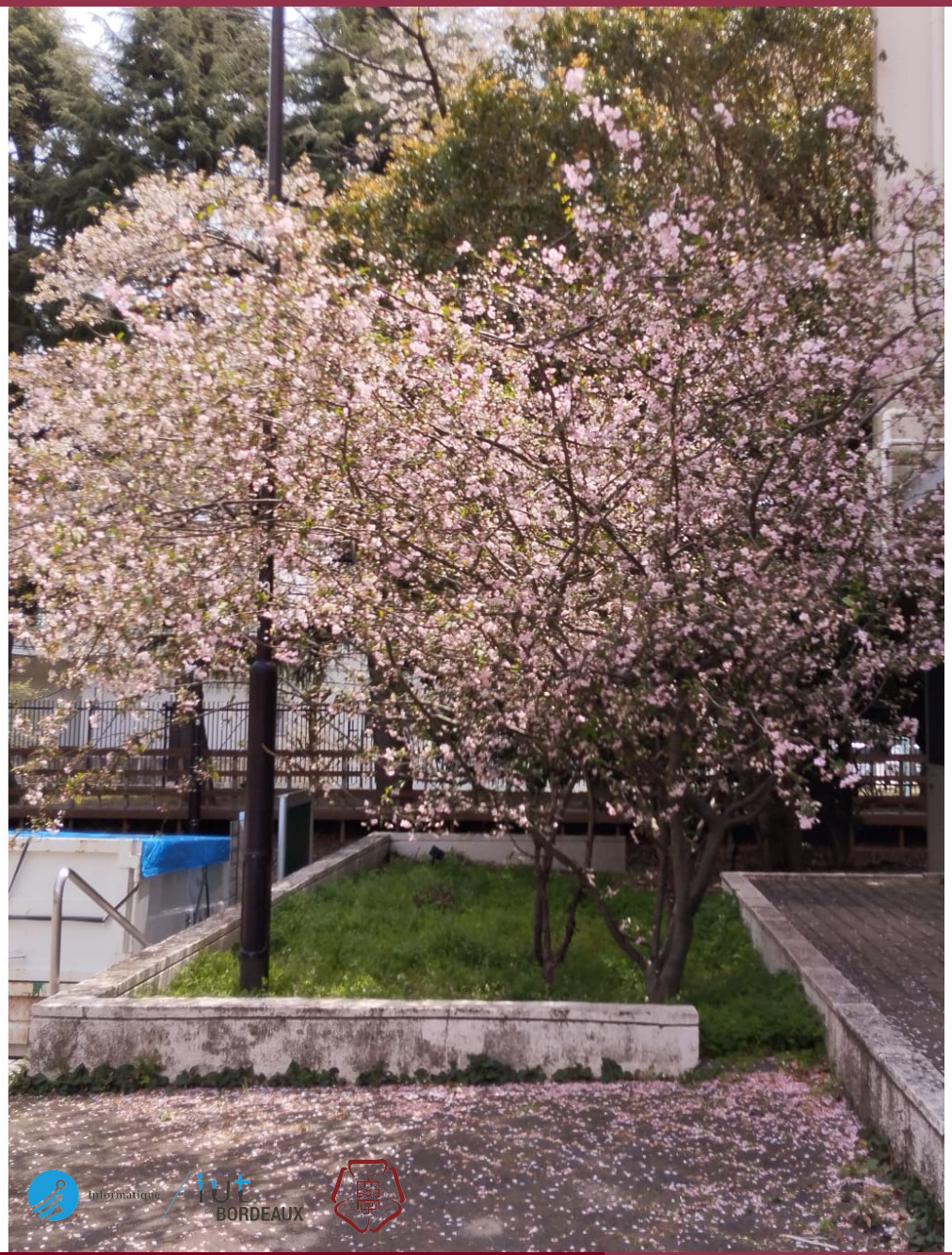
Internship
Machine Learning with Hasktorch

Candidate : GODINAT Caetano
Ochanomizu University
June 2024



contents

Structure

BekkiLab

- Focused on mathematical linguistics (fusion between Logic, Linguistics, NLP, Philosophy)
- Hosts primarily master's degree students in Information Science
- Tied to the Ochanomizu University

Ochanomizu University

- Japanese national public University
- Women University
- Multiple disciplines

contents

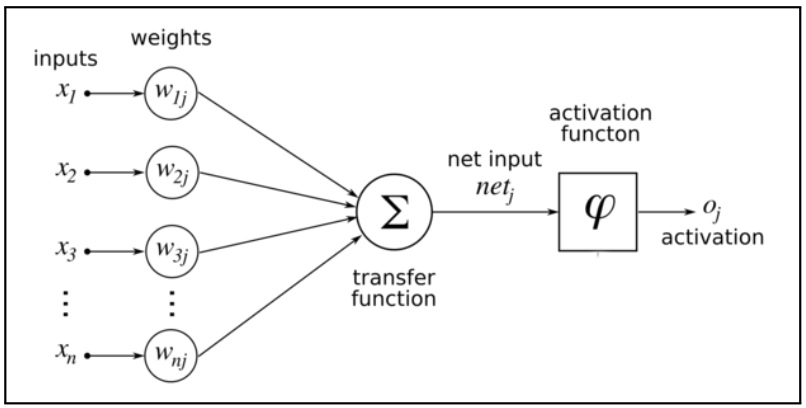
Working Schedule

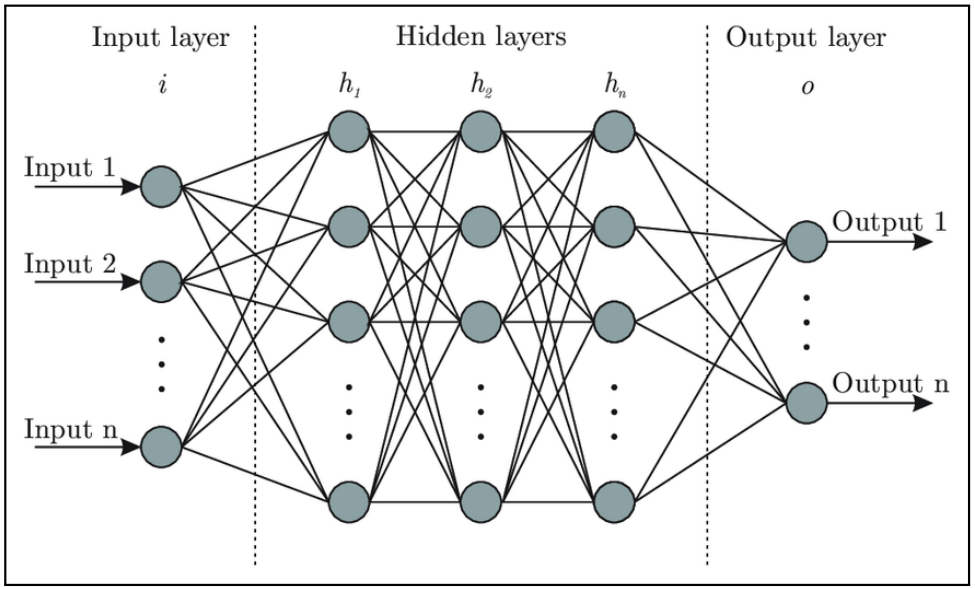
- 1st Session@April 16, 2024 Calculating Tensors in hasktorch
- 2nd Session@April 22, 2024 Linear Regression and Training
- 3rd Session@May 7, 2024 Classification, Neuron, sigmoid&softmax functions
- 4th Session@May 14, 2024 Evaluation scores
- 5th Session@May 21, 2024 Word Embeddings and Text Processing

What is it?

contents

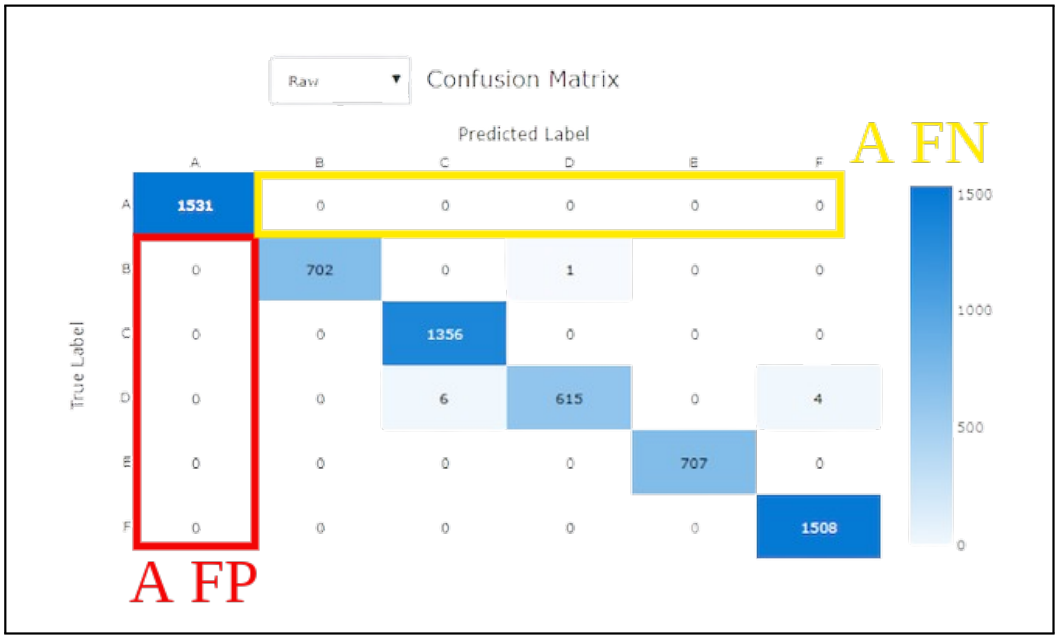
Basics of Machine Learning





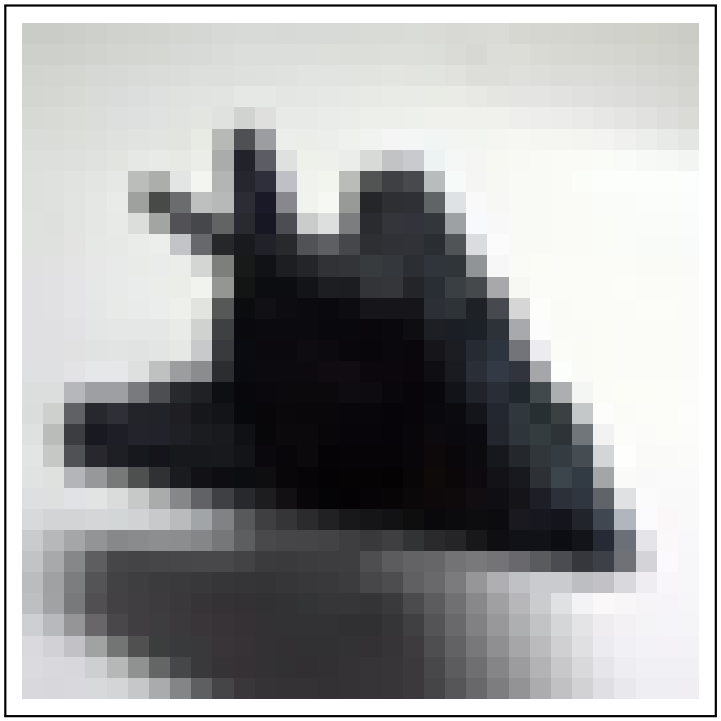
Learning

Data Analysis with F1 scores and Confusion Matrices



	precision	recall	f1-score	support
Aeroplane	0.67	0.67	0.67	3
Boat	0.25	1.00	0.40	1
Car	1.00	0.50	0.67	6
accuracy			0.60	10
macro avg	0.64	0.72	0.58	10
weighted avg	0.82	0.60	0.64	10

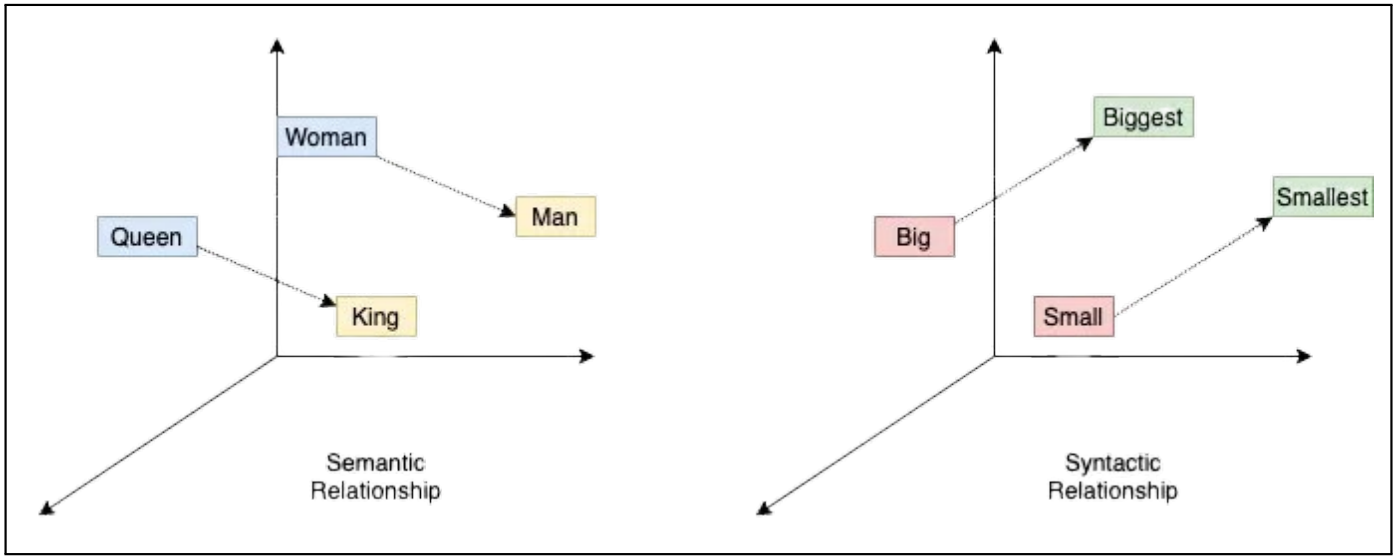
Data Analysis with F1 scores and Confusion Matrices



Epoch	Loss	Kaggle Accuracy	ValidData Accuracy	F1 Macro
50	3258.0178	-	0.4054	0.3384
550	2613.6038	-	0.4978	0.4967
1050	2321.2322	-	0.5046	0.5016
1550	2170.7373	-	0.5084	0.5063
2050	1906.2094	-	0.5088	0.5104
2550	1774.9410	-	0.5182	0.5184
3050	1658.4749	-	0.5096	0.5094
3550	1482.1902	-	0.5166	0.5176
3850	2347.8490	0.5118	0.5100	0.5071

Learning

Word to Vector




```
"mcaffee + good : [(\"mcaffee\",0.82062614),(\"really\",0.7555253),(\"install\",0.688234),(\"money\",0.68495375),(\"sh
ows\",0.65139943),(\"system\",0.64281654),(\"too\",0.6394691),(\"space\",0.6294688),(\"ads\",0.6283077),(\"transfers\"
,0.6178374)]"
```

Finding problems and Improving the library

- MLP initializer with a hasBias argument for each layer
- save and load functions for the state of the model

contents

Challenges

- Learning a new language
- Working with an unpopular technology
- Computational Power

contents

Hasktorch for Machine Learning

Pros	Cons
Best abstraction	Model manipulation is harder
Strong type system	Less Intuitive