Ham || Spam

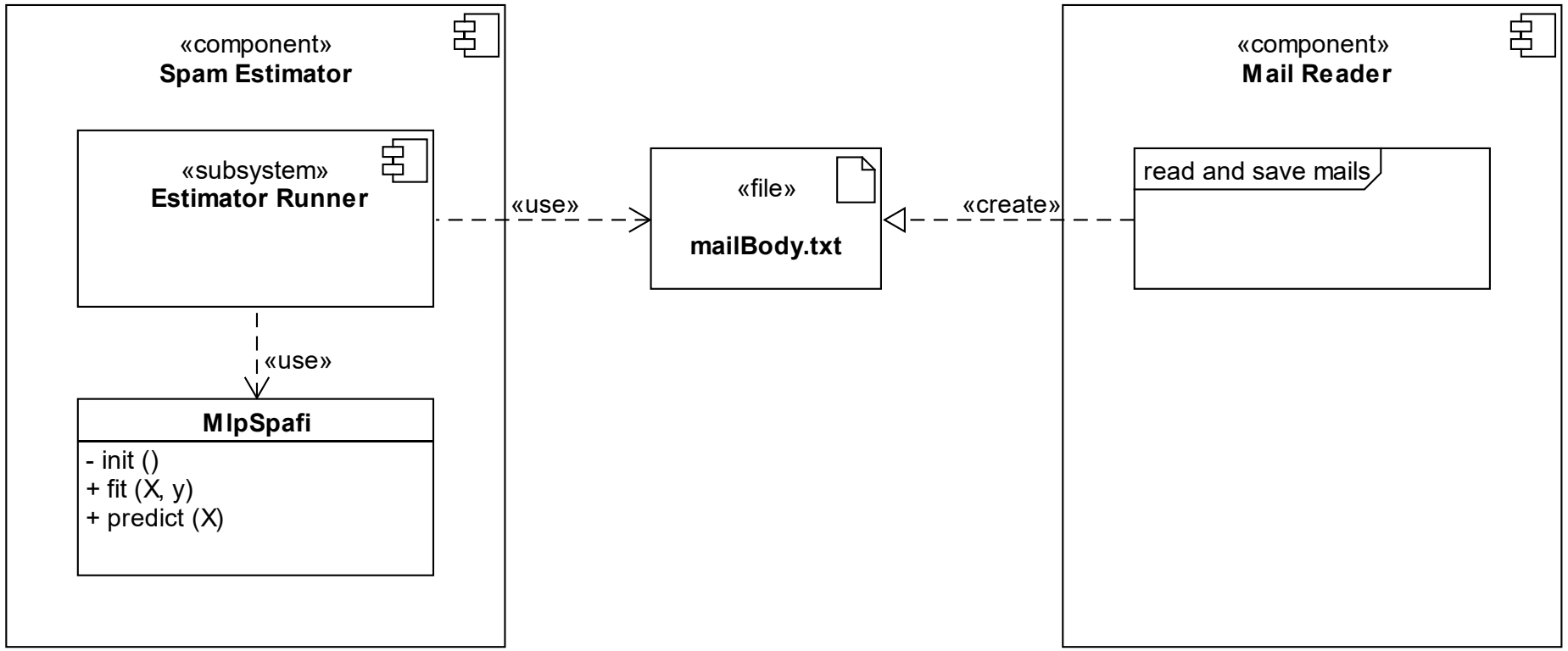
Email Spam Estimator

w/ TensorFlow and UiPath

DH

.:: Overview

Task: Identify spam using email messages



.:: Under The Hood: Spam Estimator / DE

- csv text file with 5728 emails, each labelled as spam (1) or ham (0)
- samples
  - start with ‘Subject: ‘
  - lower case text
  - contain words, numbers, non-word chars
  - min: 13 chars, max: 43952 chars
  - English only
- unbalanced, sorted dataset
  - 1368 spam, 4360 ham

Subject: contact info i will be in one of these two places - - my home : 011 91 80 3312635 my in - laws ' home : 011 91 80 5262719 you can also contact me by email at vshanbh @ yahoo . com , but it is better to call since i do not have easy access to a computer , and there may be a delay with reading email . vasant , 0

.:: Under The Hood: Spam Estimator / DP

- remove string ‘Subject: ‘
- convert text documents to a matrix of token counts

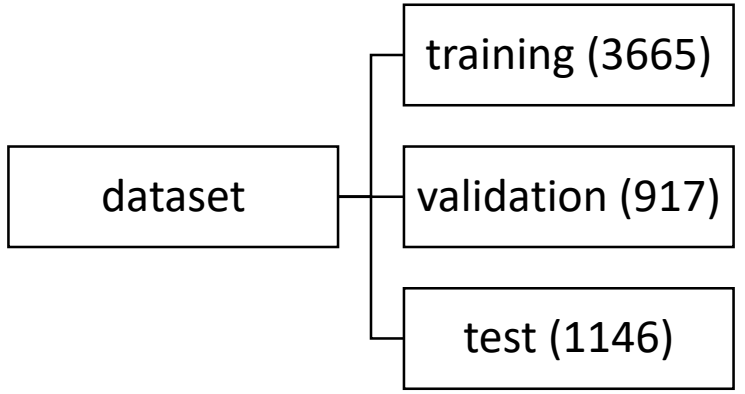
Example:

```
text_docs =
['contact info  i will be in one of
these two places’,
'but it is better to  call since i
do not have']
```

```
[[0. 1. 1. 1.]
 [1. 0. 0. 0.]]
```

```
['better',
'contact',
'info', 'places']
```

- split and shuffle dataset
  - 20% for test
  - 20% of remaining set for validation
  - balance of spam & ham persists



.:: Under The Hood: Spam Estimator / ME

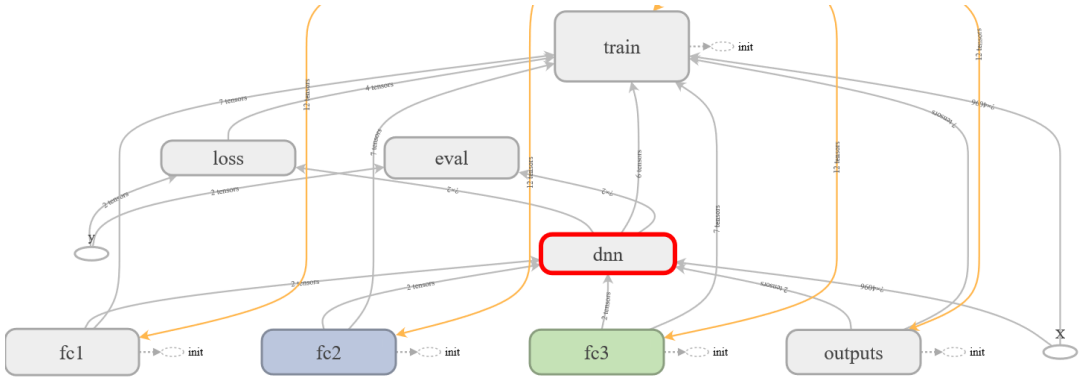
Question: Is it better to identify spam as ham or ham as spam?

Confusion Matrix:

		predicted class: spam (1)	predicted class: ham (0)
true class	spam (1)	TP	FN
	ham (0)	FP	TN

Metric to use is Precision: $PPV = \frac{TP}{TP+FP}$

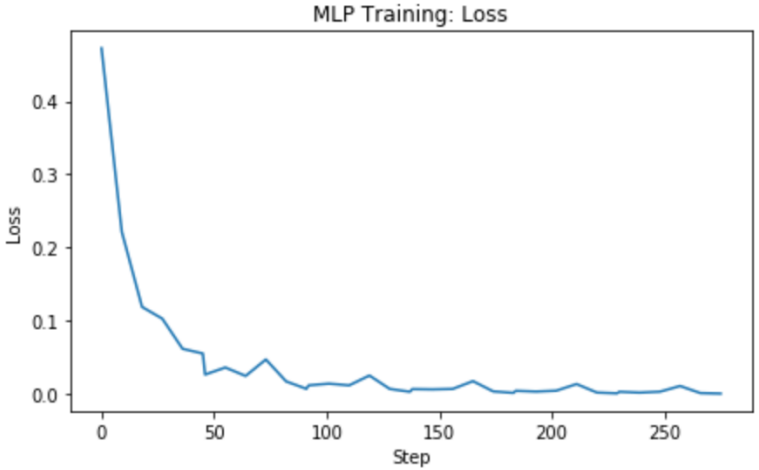
.:: Under The Hood: Spam Estimator / IM

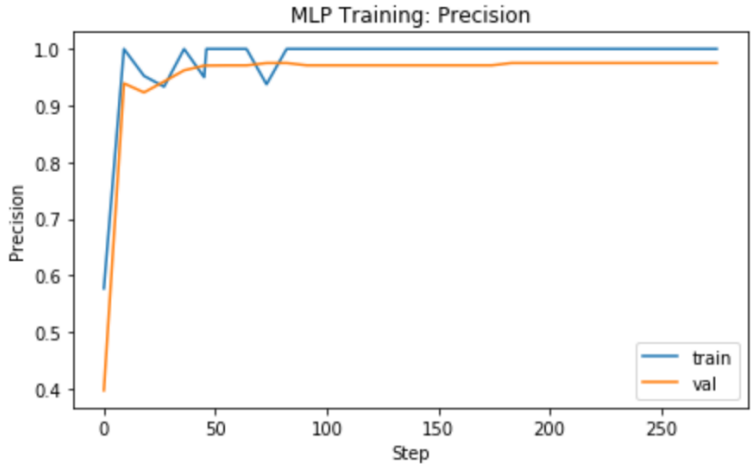


epoch # 0, batch # 46 / 46, train metric: 0.95 , val metric: 0.9704641350210971
val metric improved, model saved to ./tf_saves/run_MlpSpafi/best_acc_val.ckpt
epoch # 1, batch # 46 / 46, train metric: 1.0 , val metric: 0.9710743801652892
val metric improved, model saved to ./tf_saves/run_MlpSpafi/best_acc_val.ckpt
epoch # 2, batch # 46 / 46, train metric: 1.0 , val metric: 0.9710743801652892
epoch # 3, batch # 46 / 46, train metric: 1.0 , val metric: 0.975103734439834
val metric improved, model saved to ./tf_saves/run_MlpSpafi/best_acc_val.ckpt
epoch # 4, batch # 46 / 46, train metric: 1.0 , val metric: 0.975103734439834
epoch # 5, batch # 46 / 46, train metric: 1.0 , val metric: 0.975103734439834
early stopping after epoch 5

.:: Under The Hood: Spam Estimator / EV

Training





Scores

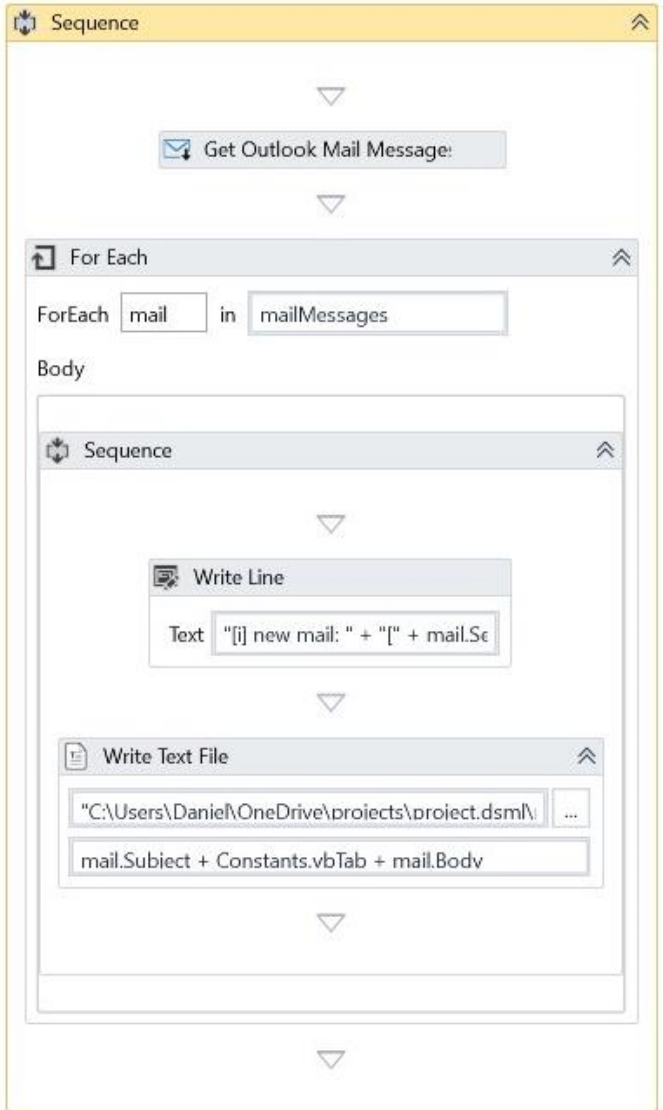
validation: 0.9751

test: 0.9893

Confusion Matrix:

		predicted class: spam (1)	predicted class: ham (0)
true class	spam (1)	280	10
true class	ham (0)	3	853

.:: Under The Hood: Mail Reader / SEQ



.:: Real-Time Demonstration

- Python script: Spam Estimator
- UiPath sequence: Mail Reader
- choose and send mails
  - http://www.antespam.co.uk/spam-resource/
  - mailto:dan.spamcheck@outlook.com

.:: Improvements

- data preprocessing
  - lemmatization / stemming
  - correlations, i.e. mail length vs spam
- implementation of estimator in UiPath
  - Python + required packages (w/o virtual environment)
  - official package ‘UiPath.Python.Activities’
- serving mode of estimator and online learning capability
- check of other mail types and languages
  - attachments
  - graphical mails

.:: Interesting Facts

- underlying MLP was prior used to recognize handwritten digits
- dataset contains mails from the Enron scandal
- less top k words worked better