

CNN vs Dense Layers

In the proposed approach, we employ a CNN to handle the code metrics. We have also tried to replace it with dense layers, and the resulting approach is presented in Fig.1. We call it Dense-based approach, and call the original one (introduced in the conference paper) as CNN-based approach.

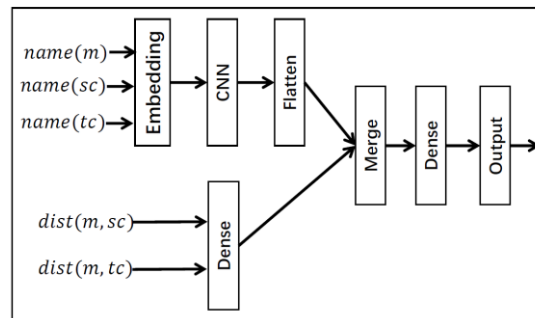


Figure 1. Replacing CNN with Dense Layer

The source code of the algorithm is presented as follows:

```
model_left = Sequential()
model_left.add(embedding\_layer)
model_left.add(Conv1D(128, 1, padding = "same", activation='tanh'))
model_left.add(Conv1D(128,1, activation='tanh'))
model_left.add(Conv1D(128, 1, activation='tanh'))
model_left.add(Flatten())
model_right = Sequential()
//we tune this dense model (e.g., number of layers and sizes of the layers) many
//times to achieve the best performance
model_right.add(Dense(16,
input_shape=(2,1), activation='tanh'))
model_right.add(Dense(16, activation='tanh'))
model_right.add(Dense(16, activation='tanh'))
model_right.add(Reshape((-1,)))
merged = Merge([model\_left, model\_right], mode='concat')
model = Sequential()

model.add(merged) # add merge
model.add(Dense(128, activation='tanh'))
model.add(Dense(2, activation='softmax'))
model.compile(loss='binary_crossentropy',
optimizer='Adadelta',
metrics=['accuracy'])
```

With the algorithm, we repeat the evaluation and results are presented on Table 1. From this table, we observe that the resulting precision and recall are very close to those of the CNN-based approach (as presented in Table 2). However, the accuracy in recommending destinations is significantly lower than of the CNN-based approach. Consequently, we keep the original CNN-based approach.

Table 1: Results of Dense-based Approach

Applications	Precision	Recall	F1	Accuracy (destination)
junit	40.59%	91.11%	56.16%	58.54%
pmd	42.86%	78.95%	55.56%	63.33%
jexcelpai	29.88%	87.50%	44.55%	34.69%
areca	48.23%	70.10%	57.14%	58.82%
freeplane	36.99%	69.73%	48.34%	48.35%
Jedit	40.93%	75.71%	53.13%	40.57%
weka	38.65%	86.39%	53.40%	42.98%
Average	38.73%	79.25%	52.03%	46.18%

Table 2: Results of CNN-based Approach

Applications	Precisi on	Recall	F1	Accuracy (destination)
junit	40.59%	91.11%	56.16%	70.73%
pmd	41.27%	68.42%	51.49%	76.92%
jexcelpai	31.90%	92.86%	47.49%	61.54%
areca	46.05%	72.16%	56.22%	71.43%
freeplane	38.09%	68.58%	48.97%	73.74%
Jedit	42.64%	78.57%	55.28%	69.09%
weka	40.05%	86.00%	54.65%	80.52%
Average	39.79%	79.27%	52.98%	74.94%