

Evernote Coding Challenge

Notebook Classifier

Chinwei Hu *

April 15, 2013

1 Introduction

This report is designated to the interview process for the machine learning internship position at Evernote, quoting the original assignment description:

Implement a Java or Python application using the Evernote service API:

`http://dev.evernote.com/documentation/cloud/`

The program should download all the notes and note meta data from the entire account. Using all the information in the account the application should:

1. Provide the names of the five most recently updated notes
2. Provide recommendations for which notebooks the notes should be placed based on the most recent 5 note's similarity to other notes in the account.

Use the algorithm you'd find most effective for determining how to classify notes by notebook. Source code should be provided. You are welcome to use 3rd party tools provided that the application can be executed solely via the command line using something like:

```
(if Python)
python classifyNotes.py
```

Background: The goal of the feature is to intelligently recommend which notebooks to place the 5 most recent notes based on how users have filed the other notes in their account. It should be a standard classification problem using the existing notes in the account as a training set then apply the model against the 5 most recent notes.

*vic@cvhu.org

2 Method

2.1 NoteClassifier1.java

References

- [1] Robertson, Stephen. "Understanding inverse document frequency: on theoretical arguments for IDF." Journal of documentation 60.5 (2004): 503-520.