

Industrial Machine Learning

Application on Automatic Email Routing &
Open Discussion on the Digital Economy

Manh NGUYEN
Cafe-Seminaire, 18th May 2019

1. Automatic Email Routing

- **Business context:**

- The client: a major global player in reinsurance
- There are several regional generic electronic mailboxes:
 - Clients and partners send emails to these generic mailboxes and expect processing and responses from the client
 - A generic mailbox is managed by a team of about 20 people who are essentially claim examiners or technical accountants
- Current configuration: to find the best person on the team to process each new email is based on both some fixed business rules and manual picking/forwarding by team members

- **Business goals:**

- Route emails automatically to the right person on the team to reduce processing time

The Proof of concept (POC)

- **Select one generic mailbox:**
 - About 20 claim examiners to manage this mailbox
 - English language mainly
- **Collecting the data on the System:**
 - 35K emails from 2015 to 2018 were collected
 - Collect the information on the person who handled each of these emails in the system, i.e the claim examiner
- **Develop machine learning solutions**

The Machine Learning blueprint

- **ML design: separate emails into:**
 - Train set: 30K emails to train models
 - Test set: 5K emails to validate/report performance
- **Preprocessing:**
 - Goals: develop appropriate ways to convert an email to a tensor (or a numerical matrix)
 - An email is an object with the Subject and the Body (and Attachment Files)
 - Use word embedding to transform texts from Subject and Body into matrices
- **Algorithm developments:**
 - Deep Learning NLP architectures
- **Performance: 92% precision on routing rate**

The Prototype

- **The machine learning solution is at the core of the prototype**
- **Advanced search for past emails:**
 - Currently, past emails once processed are stored on database and lie there forever. Users who want to review emails of past deals can only look for titles
 - The advanced search engine for emails proposes the smart search on all texts contained in emails (Subject, Author, Body and Attachment Files)
- **A further NER recognition engine to detect:**
 - Deal number, client names, claim dates, claim number etc which lie inside the email

2. Open discussion on the Digital Economy