

Author
Mathias Johansson
Approved

Date	2015-11-30	Page	1 (2)	
Document No	PRO-35218	Version	3.0	Status
Object	50048634			Approved

Thesis Work – Machine Learning and Natural Language Understanding

1 Introduction

This paper describes a thesis work that is to be carried out at TeliaSonera in Uppsala, Strandbodgatan 1. We are about 200 people responsible for TeliaSonera's contact center offering to enterprise customers. The thesis work should be performed by 1-2 students from an IT engineering, computer sciences or language technology/computer linguistics program (M.Sc.).

2 Background scenario

TeliaSonera is the market leading provider of speech enabled solutions to the enterprise market in the Nordics where we build and deliver advanced solutions based on Automatic Speech Recognition (ASR) and Natural Language Understanding (NLU). Our primary customers are large organizations with big contact centers serving their customers over the phone, online or through mobile applications.

Today's Advanced Machine Learning (ML) techniques enable increasingly sophisticated services and applications by providing means to "intelligently" process and interpret large amounts of more or less structured data. We are currently investigating how modern machine learning technology may be used for computer applications to better understand natural language. IBM's Watson and the newly open sourced TensorFlow from Google are a few examples of Neural Network based machine learning platforms that may bring new and interesting capabilities to services with a natural language interface (through speech or text).

3 Purpose

The purpose of this thesis work is to examine how a specific machine learning interface may be used to build NLU models (i.e. models that assign "meaning" to a piece of text) and how these models may be used in conjunction with speech recognition engines (i.e. technology that convert speech to text). The thesis work has three primary objectives:

1. Examine and explain how the interface provided by the chosen machine learning platform works
2. Experiment with creating various NLU models using data provided by TeliaSonera,
3. Evaluate the NLU models against a baseline created with current technology used at TeliaSonera.

If practically possible the following additional objectives may be added:

4. Run the NLU models created with a couple of speech recognition services (TeliaSonera's own platform and an online ASR service)
5. Design a confidence score algorithm to be used to assess the correctness of any given interpretation provided by the NLU model

The thesis work can be done by 1-2 students and the scope of the thesis work will be adjusted accordingly.

Company information

TeliaSonera Sverige AB
Mårbackagatan 11, SE-123 43 Farsta, SWEDEN
Registered office: Stockholm
Business ID 556430-0142, VAT No. SE556430014201

Company contact information

Visiting address: Strandbodgatan 1, Uppsala Sweden
Postal address: Box 1218, SE-751 42 Uppsala Sweden
Tel +46 8 504 55 000
Fax +46 8 5044 55 001
www.teliasonera.com

Author
Mathias Johansson
Approved

Date	2015-11-30	Page	2 (2)	
Document No	PRO-35218	Version	3.0	Status
Object	50048634			Approved

4 Suggested time plan

Following is a rough estimate of the parts the thesis work should consist of and how many weeks each part needs.

A preliminary start date would be in Jan/Feb 2016 continuing during 20 weeks.

1-3	Learn about current technology used at TeliaSonera and the ML platform currently under investigation
4-14	Experiment with NLU models using the ML platform under investigation (create models and run tests with different sets of data and configurations)
15-17	Evaluate models against existing models at TeliaSonera
18-20	Finalize thesis report

5 Applications

We are looking for students from the IT engineering, computer sciences or language technology/computer linguistic programs (M.Sc.) with an interest in a future career within IT and telecommunications. The candidate should have good experience in Java, C++ and Python programming and possess a good understanding of machine learning theory. Familiarity with speech technology is a plus. Candidates should be fluent in Swedish or English. Good communication skills and working well in a team are also important qualifications.

Your application, consisting of a personal letter and an attended courses register printout, can be sent to mathias.johansson@teliasonera.com.