



**KTH Information and  
Communication Technology**

# **Data-Efficient Deep Learning for Independent Binary Outputs**

Exploration of importance-weighted active learning, ensembling, joint training and class imbalance correction to reduce label complexity and training time in affiliate e-commerce product classification

MATTIAS ARRO

Master's Thesis at KTH Information and Communication Technology  
MSc Data Science (EIT Digital track)

Academic Examiner: Magnus Boman  
Academic Supervisor: Jim Dowling  
Industrial Supervisor: Abubakreledik Karali

2018



# Abstract

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

**Keywords:** Deep learning, machine learning, neural networks, active learning

# Referat

Denna fil ger ett avhandlingsskelett. Mer information om  
L<sup>A</sup>T<sub>E</sub>X-mallen finns i dokumentationen till paketet.

# Acknowledgment

..... London, UK, March 26, 2018  
*Mattias Arro*

# Contents

<b>Introduction</b>	<b>3</b>
Background . . . . .	3
Problem . . . . .	3
Research Questions . . . . .	3
Purpose and Goal . . . . .	3
Methodology . . . . .	3
Evaluation . . . . .	3
Work Environment . . . . .	3
Deployment Environment . . . . .	3
Ethics and Sustainability . . . . .	3
Delimitations . . . . .	3
Outline . . . . .	3
<b>Discussion</b>	<b>5</b>
<b>Conclusion</b>	<b>7</b>
<b>References</b>	<b>9</b>
<b>Declaration</b>	<b>11</b>
<b>Appendices</b>	<b>11</b>
<b>RDF</b>	<b>13</b>

# Abbreviations

LSTM Long Short Term Memory

NN Neural Network

RNN Recurrent Neural Network





# Introduction

machine learning has become successful enough to be a recurring topic in popular science and mainstream media.

## **Background**

## **Problem**

## **Research Questions**

## **Purpose and Goal**

## **Methodology**

## **Evaluation**

## **Work Environment**

## **Deployment Environment**

## **Ethics and Sustainability**

## **Delimitations**

## **Outline**



## Discussion



## Conclusion



## References





# Declaration

I hereby certify that I have written this thesis independently and have only used the specified sources and resources indicated in the bibliography.

London, UK, March 26, 2018

.....  
*Mattias Arro*



# RDF

And here is a figure

**Figure 1.** Several statements describing the same resource.

that we refer to here: 1