



Aalto University
School of Science
and Technology

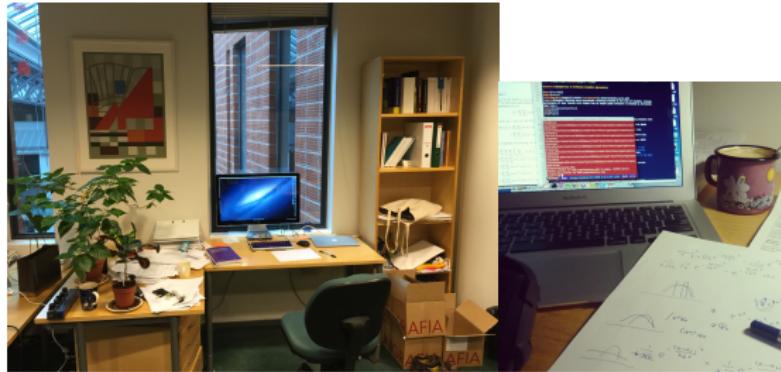
About me

Hongyu Su

Helsinki Institute for Information Technology HIIT
Department of Computer Science
Aalto University

November 5, 2015

- ▶ Hongyu Su
- ▶ Born 1984
- ▶ I am a postdoc in Helsinki Institute for Information Technology and Aalto University since 2015.5.
- ▶ Research: build advance machine learning models to solve large scale data analysis problem (research project).
- ▶ Equivalent to: continuously learning, thinking, implementing, reporting.



Educations



- ▶ Bachelor in Computer Science and Engineering, Xidian University, 2007
- ▶ Master in Bioinformatics, University of Helsinki, 2010
- ▶ Phd in Information and Computer Science, Aalto University, 2015
- ▶ Brom: *Everything comes with a price. Everything. Some things just cost more than others.*

Phd, 2011.01-2015.04, Helsinki, Finland

- ▶ The topic is **machine learning and optimization research on structured data**.
- ▶ Machine learning is to estimate outcomes (unknown) from data (known).
- ▶ Ask non-trivial machine learning questions and provide solutions.
 - ▶ Computer vision, identify object in the image.


$$(+1, +1, -1, -1, -1, +1, +1)$$

boat sea sun beach people ice land

- ▶ News articles can be assigned to multiple categories.


$$(+1, +1, -1, -1, -1, -1, -1)$$

news economics sports politics movie science art

Results

Methods and technologies that are published in TOP machine learning journal and conference.

The collage consists of several academic paper abstracts and snippets, all related to structured output learning and multi-task classification. The snippets are arranged in a grid-like structure, with some overlapping. The titles and authors of the papers include:

- Multilabel Structured Output Learning with Random Spanning Trees of Max-Margin Markov Networks** (Hongyang Su, Marko Helminen, John Shawe-Taylor)
- Multi-task Drug Bioactivity Classification with Graph Labeling Ensembles** (Hongyu Su, John Shawe-Taylor)
- Multilabel Classification through Random Graph Ensembles** (Hongyu Su, John Shawe-Taylor)
- Structured Output Prediction of Anti-cancer Drug Activity** (Hongyu Su, John Shawe-Taylor)
- Multi-label Structured Output Learning with Random Spanning Trees of Max-Margin Markov Networks** (Marko Helminen, John Shawe-Taylor, Hongyu Su)
- Structured Prediction of Network Response** (Hongyu Su, John Shawe-Taylor)
- Multi-label classification through random graph ensembles** (Hongyu Su, John Shawe-Taylor)

Each snippet provides a brief overview of the research, including the problem statement, methodology, and results. The snippets are framed by a light gray border, and the overall layout is a white background with a green horizontal bar at the bottom.

Dissertation

- ▶ My dissertation **Multilabel classification through structured output learning - methods and applications**.
- ▶ **Advanced** machine learning methods to push the boundary of multilabel classification.
- ▶ Solving many real-world **nontrivial** machine learning problems: document classification, image annotations, molecular classification, bioinformatics, social network analysis.



Work hard for 4 years and then

- Defense



- Karonkka



- Phd



Awards (name vs money)

- ▶ Chinese government awards for outstanding Phd candidate



- ▶ Some awards from Aalto university.

What did I learn from Phd?

- ▶ Strong expertise in
 - ▶ Machine learning and mathematical modelings
 - ▶ Optimization research: linear/nonlinear optimizations
 - ▶ Algorithm and data analysis / recommender system
 - ▶ Large scale data analysis
- ▶ Solid programming skills in
 - ▶ Python, Matlab, C
 - ▶ Hadoop, Spark, SQL
 - ▶ SVN, Git, Jekyll, JavaScript
 - ▶ Website and blog at www.hongyusu.com
 - ▶ GitHub at www.github.com/hongyusu
- ▶ A creative brain to solve challenging problems with modern technologies.
- ▶ An open mind that always wants to learn.
- ▶ **Be able to work hard for the long-term goal.**

Sentiment analysis

A small web application which can be built in a few days.

Twitter based sentiment analysis Twitter based sentiment analysis

Hongyu Su

Single Tweet

140 characters left

The sentiment index is 0

Hongyu Su

Single Tweet

140 characters left

The sentiment index is 0

Hashtag

50 characters left

The sentiment index is 0.434962371554

Hashtag

50 characters left

The sentiment index is 0.0519222536267

A big fan of modern information technologies

- ▶ I am curious, open, and want to learn new technologies.
- ▶ I am innovative, want to use new technologies to change daily life.
- ▶ Know new technologies very well, e.g., deep learning, big data, Kafka, Spark, IoT.
- ▶ Maintain a technical blog at www.hongyusu.com on technology innovations.
- ▶ Double blade: new are not always good, e.g., I like old-fashion mechanical keyboard.



A big fan of sports

- ▶ I enjoy competitions and aggressive sport for example basketball.



- ▶ Now I try to discover unknown part of myself
 - ▶ Downhill snowboarding.
 - ▶ Bouldering (license)
 - ▶ Paragliding
 - ▶ Open water diving (license)
- ▶ A lot of gyms.

A cat person

- ▶ Pabulo



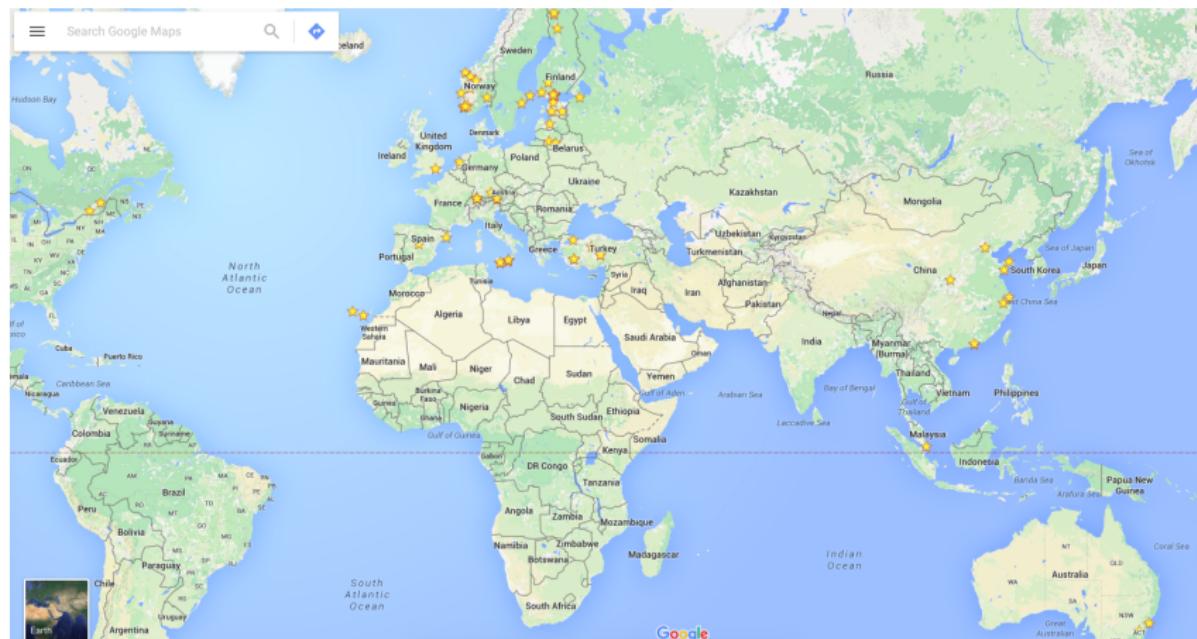
- ▶ Miu



- ▶ Cats are independent and make my life not very technical.

A hiker

I like to discover new places.



A photographer with a Flickr account

I like to keep great moments.



A bottle collector

I like to try new things.



'Stay hungry, stay foolish.' - Steve Jobs