

SUMMARY

NLP specialist with hands in data scraping and data mining, A Python man and a developer of web apps that are mostly run with Machine Learning algorithms. Perfectionist DiSC Pattern and an enthusiast on creating new things and working on challenging projects.

RESEARCH INTERESTS

Natural Language Processing, Data Mining, Machine Learning, Web mining, Information extraction, Social Network Analysis

EDUCATION

2016-2018, **Sharif University of Technology**, Master of Science (M.Sc.), Computational Linguistics.

Diss: "Deep Neural Network Language Modelling" Dr. Hossein Sameti

2012-2016, **Imam Khomeini International University**, Bachelor of Arts (B.A.), Teaching English as a Foreign Language (TEFL)

PUBLICATIONS

Thesaurus based query expansion for Persian, Working Paper

Detecting Hoax Content on Wikipedia: A Naïve Bayesian Approach, Adel Rahimi, 5th international conference on artificial intelligence, computer science & information technology 2017, Penang, Malaysia.

Improvement of English to Persian Machine Translation via N-grams of Part-of-Speech tags, Adel Rahimi, Proceedings of 3rd regional conference on new achievements in electrical and computer engineering 2017, Mazandaran, Iran

The phonetics and the phonology of final boundary tone in Northern Kurdish, Adel Rahimi, Lingbuzz/002902, 2015.

A hybrid stemming algorithm for Persian, Adel Rahimi, ArXiv: 1507.03077, 2015.

Lexical Bundles in Computational Linguistics' academic literature, Adel Rahimi, arXiv:1603.02905, 2015.

PROFESSIONAL EXPERIENCE

Snappfood.com (subsidiary of Rocket Internet), Business Intelligence Expert **Jun 2017-now**

Developed NLP tools for Elasticsearch: Stemming, Named Entity Recognition and Query Understanding. Implemented OLAP Datawarehouse based on SSAS. Created several reports based on data mining techniques such as Sales and user prediction, server load prediction, and future KPI prediction based on different libraries like: FB Prophet, Scikit-learn. Built KPI dashboards and reports on Power BI and SSRS. Developed Bayesian rating system. Developed Customer and Vendor segmentation based on clustering algorithms.

Sharif Speech and Language Processing Lab, Researcher **Feb 2016-now**

Head of the team for building Persian programmable chat-bot; First programmable chat-bot for Persian. Tasks in this project included: Developing ERP, Developing FSA, managing issues and tasks, developing Naïve Bayesian based machine learning algorithm for the bot. revising and analyzing codes and issues.

Chartex.ir, Backend Developer, Web scraping specialist **Dec 2016-Apr 2017**

Developed and mapped a database of 300K+ hotels and developed API handler for hotel room reservation, developed algorithm for transliterating Persian names from Latin script to Arabic, Developed and implemented crawler caching system to increase search speed, Developed and implemented Data-driven prediction models for ticket reservation

Takvaj.ir (Sharif NLP Group), Back-End Developer **2016-now**

Developed Persian tokenizer and normalizer (available at takvaj.herokuapp.com)

Nested.me, NLP Specialist and Backend Developer

Sep-Nov 2016

Developed product-info crawler and price crawler, Developed text summarization project

Teaching Assistantship, Dr Parvaneh Khosravizadeh

Mar 2016

International Association of Engineers (IAENG), Member

2013-now

Societies for Artificial Intelligence, Data Mining, Computer Science, and Software Engineering

SKILLS AND EXPERTISE

Programming Languages: Python (NLTK, Bokeh, Matplotlib, Scipy, Flask, Selenium, Requests, Beautifulsoup, Scrapy, Django, scikit-learn), R: NLP, Text mining (TM), HTML CSS (Responsive Designs)

Databases: Elasticsearch, MongoDB (PyMongo) and other NoSQL databases, MySQL and similar SQL databases

Deployment tools: Docker, Vagrant, Travis Ci, Jenkins

Big Data: Beginner familiarity with the Hadoop framework and MapReduce algorithm

Data Visualization: Chart.js, Power BI, tableau, RDF, OWL

Project Management: Quality Assurance and Software testing, Familiar with CMMI, familiar with SDLC, Agile method, Jira

Other: Git (Github, Gitlab, and Bitbucket), Familiar with Machine Learning approaches, Rapidminer, SPSS (Data Mining), Weka, Orange, Apache OpenNLP, Object-oriented designs and MVC Designs, Linux and Mac (Primary), Wordpress, familiar with cloud tools such as Heroku and Amazon S3

CODES AND PROJECTS

ArticleIE, built an Information extraction bundle based on python to extract latest trend in Journal Articles

Persian Query Understanding, built a query understanding algorithm and it's data based on thesaurus

PersianChatbot, Persian Chatbot that uses naïve Bayesian Algorithm to understand user intent (Supervised by: Dr Bahram Vazirnejad)

TextSlide, Automatic generation of presentation slides from natural language texts, backed by NLP algorithms and written in Python, released under GNU, available at: textslide.herokuapp.com

Kurmanji-Stemmer: The second release. Zenodo. DOI: 10.5281/zenodo.29605

Balaxan corpus of Kurmanji, LINDAT/CLARIN digital library at Institute of Formal and Applied Linguistics, Charles University in Prague, HDL:11372/LRT-1531

Kurmanji Wikipedia Corpus, DOI:10.7910/DVN/OHWSUI, Harvard Dataverse Network, V1

Corpus of World National Anthems, DOI:10.7910/DVN/PZG8TH, Harvard Dataverse Network, V1

Dakhil wordlist for Persian vocabulary, DOI:10.7910/DVN/MJBHLN, Harvard Dataverse, V1

Corpus of Computational Linguistics' Academic Literature, DOI:10.7910/DVN/YHHTCI, Harvard Dataverse, V2

TALKS AND SESSIONS

Summer 2017, 5th international conference on artificial intelligence, computer science & information technology, oral presentation on “**Detecting Hoax Content on Wikipedia: A Naïve Bayesian Approach**”

Spring 2017, 3rd regional conference on new achievements in electrical and computer engineering, Presentation on “**Improvement of English to Persian Machine Translation via N-grams of Part-of-Speech tags**”

Spring 2017, Sharif University of Technology, Session titled “**Big Data + Sentiment Analysis = Awesome**”, An introduction on Big data using NLP tools in MapReduce algorithm.

Fall 2017, Sharif University of Technology, Session titled “**Persian Intonation**”.

Fall 2017, Sharif University of Technology, Session titled “**X-bar Theory**”.

LANGUAGES

English Fluent, TOEFL iBT 96 (2015)

Kurmanji Native Speaker

Persian Native Speaker

High Arabic Reading knowledge

AWARDS AND HONORS

Awarded merit based 2 years free education for Master's degree

Ranked 42 in the nationwide university entrance exam for Masters i.e. Konkour (top 99.998%)

Awarded merit based 4 years free education for Bachelor's degree

Ranked 900 in the nationwide university entrance exam for Bachelors i.e. Konkour (top 97%)