800 Sunset Ave. APT#2,
Manhattan, KS, 66502

Krutarth Patel

kipatel@ksu.edu
(940) 999-4444

http://www.linkedin.com/in/krutarth-patel-50805b97
https://patelkrutarth.github.io/

Education

Ph. D. candidate, Computer Science (4.0 GPA)	Kansas State University, Manhattan, KS	Expected May 2021
M.S., Computer Science (4.0 GPA)	University of North Texas, Denton, TX	Dec. 2017
B. Tech., Computer Engineering	Dharmsinh Desai University, Gujarat, India	May 2014

Interests & Skills

Interests	: Deep Learning, Machine Learning, Information Retrieval, Natural Language Processing, Data Mining
Programming	: Python, Java, HTML, CSS, JavaScript, C++, C, Matlab
Frameworks and Tools	: TensorFlow, PyTorch, Theano, Keras, scikit-learn, nltk, pandas, numpy, BERT, Word2Vec, FastText, SciBERT, Weka, WordNet, Mallet, PostgreSQL, SQL, Django, Hadoop, PySpark, Amazon Web Services (AWS: EC2, S3)

Work Experience

PhD Intern- Machine Learning, enotice Jun. 2019 - Aug. 2019

Technologies: Python, nltk, scikit-learn, regular expressions, Django, PostgreSQL, HTML, CSS

- Designed and developed machine learning algorithms for classification and analysis of the company's proprietary dataset of Public Notices containing unstructured text data.
- Designed scrapers to scrape public notices and store those notices into a PostgreSQL database using Django framework.

Graduate Research Assistant Mar. 2015 – Present

Technologies: Python, Java, TensorFlow, PyTorch, Theano, Keras, scikit-learn, nltk, pandas, numpy, Word2Vec, FastText, BERT, SciBERT, WordNet, Mallet, regular expressions, Amazon Web Services (AWS: EC2, S3), Weka, SQL, MySQL, Bing search API

PDFMEF: Multi-Entity Knowledge Extraction Framework for Scholarly Documents (https://github.com/SeerLabs/pdfmef)

- Integrated a module for the keyphrase extraction task.
- Integrated a figure extraction module to use the pdffigures2 and fixed a bug to handle the XML output generated from the GROBID tool.

Keyphrase extraction

- Formulated keyphrase extraction task as a sequence labeling using Conditional Random Fields (CRFs) and Bi-LSTM CRF.
- Trained word Glove and Word2Vec embeddings for using them as features along with other document specific features.
- Implemented an unsupervised approach that makes use of the theme of a given document as well as the positional information.
- Used multiple word embeddings based on Word2Vec, FastText, BERT, and SciBERT to get the theme vector of a document.
- KPRank (our unsupervised approach): **https://github.com/PatelKrutarth/KPRank**

Identifying documents relevant to a given collection from Web Archives

- Implemented deep learning classifiers, structural features classifiers, and BoW classifiers for a document classification task.
- Applied feature selection methods the dimensionality of the features.
- Implemented dynamic model selection algorithms to dynamically select the best classifier on-fly. Also used competence learning for dynamically selecting the most competent classifier.

Homepage Classification and Discovery

- Designed deep learning-based classifiers (CNN, RNN, and LSTM).
- Used the deep learning-based classifier in the researcher homepage discovery framework.
- Used self-learning to train a classifier by utilizing unlabeled data and to reduce the human effort for labeling the data.

Identifying research articles from the collection of documents

- Implemented deep learning architectures, structural features, and a co-training approach for identifying research articles.
- Used those classifiers in the search-driven framework for collecting research articles.
- Adapted those classifiers for collecting relevant documents from web archiving collections.

Software Engineering Intern, ISRO (Indian Space Research Org.), India Dec. 2013 - Apr. 2014

Technologies: C++

- Modified the implementation of Layer-2 LAPDm of GSM stack for the disaster management.

Academic Projects

Web Search Engine, University of North Texas Mar. 2016 - Apr. 2016

Technologies: Java, Weka, regular expressions, SQL, MySQL

- Created and deployed a Web spider and integrated it with an information retrieval system based on the vector-space model.
- Implemented additional user feedback to update relevancy ranking algorithm.

Part of Speech Tagger for Tweeter Sentences, University of North Texas Nov. 2016 - Dec. 2016

Technologies: Python, nltk, numpy, regular expressions

- Used bigram and trigram Hidden Markov Model along with morphological rules.
- Used Viterbi algorithm for dynamic programming.

Eagle's Eye (Bus tracking android application for the UNT buses), University of North Texas Apr. 2017 - May 2017

Technologies: Android, Java, WebSocket API, Google Maps API, Directions API, Retrofit

- Buses currently running on the selected route are highlighted on the screen with a live update of its location. The nearest bus stop showed to the user. User can comment on the bus. Buses are colored red or green depending on bus is full or not.
- Implemented a server to retain comments and the live bus position.

Teaching Experience

Graduate Teaching Assistant, Kansas State University, USA Aug. 2020 – May. 2021

- Subject: Deep Learning (08/2020 to 12/2020), Operating Systems (01/2021 to 05/2021)

Graduate Teaching Assistant, Kansas State University, USA Aug. 2019 - May. 2020

- Subject: Deep Learning (08/2019 to 12/2019), Big Data (08/2019 to 12/2019), Operating Systems (01/2020 to 05/2020)

Graduate Teaching Assistant, Kansas State University, USA Aug. 2017 - Dec. 2017

- Subject: Information Retrieval

Teaching Assistant, University of North Texas, USA Jan. 2017 - May 2017

- Conducted two labs for the course Computer Science I, which included basic C/C++ programming.

Guest Lectures

- Gave several lectures in graduate courses on Machine Learning and Information Retrieval (Dr. Cornelia Caragea)

Achievements & Awards

- The Innovative Application Award in the AAAI/IAAI 2021 conference.
- The Best Student Paper Award in the JCDL 2020 conference.
- Received the Toulouse Graduate School (TGS) Tuition Benefit Scholarship 2015-2017.
- Received the Dean Tuition Scholarship Summer 2015.
- Graduated 2nd ranked student in the Computer Engineering department and appeared among top 2% of all students in the entire Dharmsinh Desai University.

Leadership & Activities

- Committee member of the CS-GSC (Computer Science-Graduate Student Council) group at K-State.
- Represented the Computer Science Department at UNT in the Graduate Study Expo (September 2015).
- Represented the Computer Science Department at UNT in the event for new international students (September 2015)
- Presented a research study on a navigation game for children at the Fort Worth Museum of Science and History (December 2015)
- Administered events committee at Felicific 2013 - The Cult & Tech Fest of the DDU**,** Dharmsinh Desai University
- Food committee member in the BAPS Swaminarayan Chhatralaya/hostel (2010-2013)

Publications

- **Krutarth Patel**, and Cornelia Caragea. "Exploring Word Embeddings in CRF-based Keyphrase Extraction from Research Papers" K-CAP, 2019.
- Mark Phillips, Cornelia Caragea, **Krutarth Patel**, and Nathan Fox. "Leveraging Machine Learning to Extract Content-Rich Publications from Web Archives." WAC, 2019.
- Mark Phillips, Cornelia Caragea, **Krutarth Patel**, and Nathan Fox. "Increasing Access to Content-Rich Publications from Web Archives with Machine Learning Models." TCDL, 2019.
- **Krutarth Patel**, Cornelia Caragea, and Mark E. Phillips. "Dynamic Classification in Web Archiving Collections" LREC, 2020.
- **Krutarth Patel**, Cornelia Caragea, Mark E. Phillips and Nathaniel T. Fox. "Identifying Documents In-Scope of a Collection from Web Archives" JCDL, 2020. (Won the Best Student Paper Award)
- **Krutarth Patel**, Cornelia Caragea, Jian Wu and C. Lee Giles. "Keyphrase Extraction in Scholarly Digital Libraries Search Engine" ICWS, 2020.
- **Krutarth Patel**, Cornelia Caragea, and Sujatha Das Gollapalli. "On the Use of Web Search to Improve Scientific Collections" SDP@EMNLP, 2020
- **Krutarth Patel**, Cornelia Caragea, Doina Caragea, and C. Lee Giles. "Search Driven Author Homepage Discovery using deep learning architecture" IAAI, 2021. (Won the Innovative Application Award)
- **Krutarth Patel**, and Cornelia Caragea. "Exploiting Position and Contextual Word Embeddings for Keyphrase Extraction from Scientific Papers" EACL, 2021.
- **Krutarth Patel**, Mark E. Phillips, and Cornelia Caragea. "Using Competency Algorithm to Find Relevant Documents from Web Archives" to be submitted to JCDL, 2021.