Abhay Kashyap

Data Scientist, Software Engineer

443-535-7524
abhay1@umbc.edu
abhaykashyap.com

Summary

Research experience in Machine Learning and Text Processing;
Engineering experience in full stack web development, responsive design and interactive visualizations.

Technical Skills

Programming	Python, JavaScript, MATLAB/Octave, C
Web Frontend	HTML5, CSS3, Bootstrap, Less, Sass, D3
Web Backend	Django, Nginx, PostgreSQL
Tools/Software	NLTK, Stanford CoreNLP, Scikit, Weka, Gensim, TensorFlow, Inkscape

Education

Aug '11 - Present
PhD in **Computer Science**, University of Maryland at Baltimore County
Adviser: Dr. Tim Finin; **GPA:** 3.67
Thesis: Lyrics Augmented Multi-modal Music Recommendation

Aug '09 - June '11
MS in **Computer Science**, State University of New York at Buffalo
Adviser: Dr. Venu Govindaraju; **GPA:** 3.67
Project: Facial behavior as a soft biometric

Aug '05 - July '09
BE in **Computer Science**, Visvesvaraya Technological University, Bangalore

Work

Sept '16 - Present
Sept '15 - Dec '15
Course Instructor, University of Maryland at Baltimore County
Course: Introduction to Artificial Intelligence (CMSC 471)

Summer '15, '14
Data Scientist Intern, @WalmartLabs, San Bruno, California
Manager: Dr. Pranam Kolari

'15
Built distributed representations for e-commerce items using word2vec on item click stream data to improve item recommendations.

'14
Enhanced the existing taxonomy of e-commerce items using a combination of graph based clustering on co-occurrence data and standard machine learning classification methods on attributes.

Jun '13 - May '15
Research Assistant, University of Maryland at Baltimore County

Aug '11 - Dec '12
Teaching Assistant, University of Maryland at Baltimore County
Courses: Computer Architecture, Cryptology

Key Projects

May '14 - Present

Rapalytics (rapalytics.com)

Featured on Vox, Metro, Exclaim, Attn: etc.

Description

Artists express thoughts and emotions through lyrics, especially in genres like Rap, but contemporary collaborative-based recommenders largely neglect them. Rapalytics fills this gap by computing lyrical similarity between songs and artists for music recommendation, personalization and discovery. In its pilot launch, Rapalytics let users explore features central to lyric composition such as rhyme density, simile usage and profanity levels through responsive, interactive visualizations. Upcoming release will feature lyrics augmented music recommendation & integration with streaming services like Spotify.

Responsibilities

Solo project; designed & developed the concept & product in entirety.

Technology, Tools & Resources

Python, Django, PostgreSQL, Nginx; HTML5, CSS3, Bootstrap, Less, JavaScript, JQuery, Ajax, D3; Inkscape; NLTK, Stanford coreNLP, gensim, word2vec; MusicBrainz, Echonest, Last.fm

Related Talks

SF Text Meetup, SF Data Visualization Meetup, Data Science MD Meetup

---

Sept '15 - Present

SafePod (safepodapp.org)

Description

SafePod is an android app aimed to provide a safe way to anonymously share traumatic experiences that include domestic violence, sexual assault, depression etc. with the goal of providing a platform to find support and help recovery. Kickstarted during the *Baltimore hackathon 2015*, it is envisioned as an open effort to promote health and safety.

Responsibilities

Product design and management, UX, Web development

Jan '13 - Mar '13

Semantic Textual Similarity (STS)

Description

The objective of STS, a SemEval workshop task, is to build system that measure the degree of semantic equivalence between sentences. The developed system used Latent Semantic Analysis (LSA) and simple term alignments to compute sentence similarity. Our team ranked 1st among 35 participants in 2013 and 2nd out of 15 teams in 2014.

Responsibilities

Extended word similarity to handle longer text pieces by extracting several features and training Support Vector Regression models to predict scores.

Technology, Tools

Python, Stanford coreNLP, DBpedia, NLTK, LIBSVM

Aug '10 - May '11

Facial behavior as a Soft Biometric (Masters Project)

Video segments containing facial expressions of a set of individuals were annotated using Facial Action Coding System (FACS). These features were then used to uniquely identify individuals with a significantly higher than

random accuracy which indicated the presence of facial behavioral signatures.

Publications

2016 Das, Prajit Kumar, Abhay Kashyap, Gurpreet Singh, Cynthia Matuszek, Tim Finin, and Anupam Joshi. "Semantic knowledge and privacy in the physical web." In Proceedings of the 4th Workshop on Society, Privacy and the Semantic Web-Policy and Technology (PrivOn 2016) co-located with the 15th International Semantic Web Conference (ISWC 2016). 2016.

2015 Kashyap, Abhay, Lushan Han, Roberto Yus, Jennifer Sleeman, Taneeya Satyapanich, Sunil Gandhi, and Tim Finin. "Robust semantic text similarity using LSA, machine learning, and linguistic resources." *Language Resources and Evaluation*: 1-37.

2014 Kashyap, Abhay, Lushan Han, Roberto Yus, Jennifer Sleeman, Taneeya Satyapanich, Sunil Gandhi, and Tim Finin. "Meerkat mafia: Multilingual and cross-level semantic textual similarity systems." In *Proceedings of the 8th International Workshop on Semantic Evaluation*, pp. 416-423. Association for Computational Linguistics, 2014.

2013 Han, Lushan, Abhay Kashyap, Tim Finin, James Mayfield, and Jonathan Weese. "UMBC EBIQUITY-CORE: Semantic textual similarity systems." *Atlanta, Georgia, USA* 44 (2013).

2012 Kashyap, Abhay L., Sergey Tulyakov, and Venu Govindaraju. "Facial behavior as a soft biometric." In *Biometrics (ICB), 2012 5th IAPR International Conference on*, pp. 147-151. IEEE, 2012.

Activities

Participated in the Baltimore Hackathon, 2015; and launched SafePod App

Participated in hackUMBC, 2016 and was in the top 10 among 70+ teams.

Part of the Program Committee at Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL), 2013, 2014 & 2015.

Volunteered at R.A.R.E, a non-profit organization for the rescue and rehabilitation of reptiles.

Volunteered at CWS (Center for Wildlife Studies) as a line transect survey assistant