TANVI SAHAY

Brittany Manor Drive, Amherst, MA - 01002
tsahay@cs.umass.edu | | | tsahay@people.cs.umass

Education

May 2018	**University of Massachusetts Amherst, MA**, Master of Science, Computer Science	4.0
June 2016	**Birla Institute of Technology, India**, Bachelor of Engineering	3.0

Skills

Python, C, tensorflow, scikit-learn, Stanford CoreNLP, Keras, SQL, MATLAB, R, Git, Javascript

Relevant Subjects

Spring 2017	Data Visualization and Exploration	A
	Database Design and Implementation	A
Fall 2016	Machine Learning	A
	Introduction to Natural Language Processing	A

Work Experience

June 2017 **R&D Intern, Lexalytics Inc.**

Feb 2017 -May 2017 **Data Science Independent Study, Lexalytics Inc.**

- Concept/Theme Rollup of Text
- Extracting key phrases from a set of sentences and representing them as real-valued vectors to allow clustering of similar phrases into a single group.

Selected Research Projects

Feb 2017 -May 2017 **Schema Matching using Machine Learning**

- Engineered custom features to represent schema names and employed unsupervised machine learning algorithms to cluster similar schema names
- Performed within-cluster one-to-one matching using edit distance and introduced the idea of domain-based global dictionary for the purpose of one-to-many schema matching

Oct 2016 -Dec 2016 **Sentence Generation using Fan Theories**

- Employed CoreNLP package for tokenization, relation extraction, PoS tagging and Named Entity Recognition on a database of Fan Theories of Game of Thrones
- Performed noise removal using OpenIE and used Bigram, HMM and character-level LSTMs for sentence generation.
- Analyzed results based on overall coherence, general fluency and information content with and without considering domain knowledge.

Oct 2016 -Dec 2016 **Architecture Classification for Indian Monuments using ORB features**

- Extracted ORB features of monument images and performed architecture classification using KNN, Logistic Regression, SVM and Random Forests.
- Compared image-wise classification using different supervised techniques with descriptor-wise classification using KNNs.

Aug 2015 -May 2016 **A.T.O.M - A Tool for Music transcription using machine learning and image processing**

- Recorded a raw database for audio-visual recognition of isolated guitar notes for over 50 users.
- Extracted MFCC features of guitar notes and applied SVM for note recognition.
- Used position markers and relative finger position to confirm identity of the note.

Selected Publications

- A. Aggarwal, R. Kumar, T. Sahay and M. Chandra, "GuiTones-I: An Audio-Visual Database of Monophonic Guitar Tones", *TENCON 2016 - 2016 IEEE region 10 conference*, Singapore.
- A. Aggarwal, T. Sahay, A. Bansal and M. Chandra, "Grid search analysis of nu-SVC for text-dependent speaker-identification," 2015 Annual IEEE India Conference (INDICON), New Delhi, 2015. *Best Paper Award*