

# Sachi Angle

sva22@cornell.edu | (857) 361 – 8852 | [github.com/sach211](https://github.com/sach211)

<b>EDUCATION</b>	<b>Cornell Tech at Cornell University</b>	<b>New York, NY</b>
May 2021	M.S. in Information Systems and Applied Information Sciences, Concentration in Connective Media <i>Courses: Natural Language Processing, Applied Machine Learning, Deep Learning, Tech Media and Democracy</i>	
May 2018	<b>Manipal Institute of Technology</b> B.Tech. in Computer Science and Engineering <i>Courses: Natural Language Processing, Machine Learning, Artificial Intelligence Design and Analysis of Algorithms (Teaching Asst), Data Structures</i>	<b>Karnataka, India</b>
<b>RESEARCH EXPERIENCE</b>	<b>MT-NLP Lab at the International Institute of Information Technology</b>	<b>Hyderabad, India</b>
Jan 18 – May 18	<i>Research Intern</i> <ul style="list-style-type: none"><li>Built a command line tool for Part of Speech Tagging of Konkani (A data scarce language) with 77% accuracy using Semi-Supervised Machine Learning Techniques like Active Learning and Self Training</li><li>Built an Automated Error Correction tool for POS tags of Hindi Data, in order to ensure that higher quality data is available for use. Achieved an accuracy of 94.02% with a Neural Network that employed the probabilistic predictions of Support Vector Machines, Long Short Term Memory Networks, and Conditional Random Fields to make its predictions</li></ul>	
2016 – 2017	<b>Manipal Institute of Technology</b> <i>Research Intern</i> <ul style="list-style-type: none"><li>Built a Morphological Analyzer for Kannada, an agglutinative language spoken in Karnataka, India, using Support Vector Machines for classification and achieved an accuracy rate of 85%</li><li>The morphological components of the language provide valuable information regarding the properties of the word (e.g. Plurality, Gender) which can be used as data for other complex ML tasks</li></ul>	<b>Karnataka, India</b>
<b>PROFESSIONAL EXPERIENCE</b>	<b>Microsoft</b>	<b>Bangalore, India</b>
2018 - 2019	<i>Software Developer at Microsoft IDC</i> <ul style="list-style-type: none"><li>Worked on the pipeline to efficiently and compliantly extract raw Office 365 data, in parallel, as structured schematized data available for consumption by partner ML teams.</li><li>Filed for a patent in the scope of using NLP to improve user experience</li></ul>	
May 17 – July 17	<b>Adwyze</b> <i>Data Analyst Intern</i> <ul style="list-style-type: none"><li>Adwyze is a start-up providing Marketing insights to their customers. I worked on a natural language interface to enable clients to build dashboards and analyse their sales data by typing the commands in natural languages, as opposed to programmatic query languages.</li></ul>	<b>Bangalore, India</b>
<b>PROJECTS</b>	<b>Soccer Summary Generator</b> Built a python command line tool which generates a textual summary of a video of a football match using RNNs and CNNs	
<b>SKILLS</b>	<b>Programming: Python, C#, C++</b> <b>Libraries: Pandas, Scikit-learn, Numpy</b>	
<b>ACHIEVEMENTS</b>	<ul style="list-style-type: none"><li>Cornell Tech Merit Scholarship Award</li><li>Publication: “Automated Error Correction and Validation for POS Tagging of Hindi”, ACL, PACLIC 2018 (<a href="https://www.aclweb.org/anthology/Y18-1002/">https://www.aclweb.org/anthology/Y18-1002/</a>)</li><li>Publication: “Kannada morpheme segmentation using machine learning”, IJET 2018 (<a href="https://www.sciencepubco.com/index.php/ijet/article/view/13395">https://www.sciencepubco.com/index.php/ijet/article/view/13395</a>)</li><li>First runner-up at Microsoft Bangalore for the annual hackathon</li></ul>	
<b>LEADERSHIP</b>	<ul style="list-style-type: none"><li>Member of Google’s Women Techmakers</li><li>Teach Code for Good, a volunteer program through which underprivileged children are taught how to code in python - Board Member and Mentor.</li></ul>	