

Pankaj Bhambhani

Masters Student at UMass Amherst, 3 years of Professional Experience. Seeking Machine Learning, NLP, Data Science, or Dev Internships.

✉ pbhambhani@cs.umass.edu ☎ 413-230-6252 🌐 [pankajb64.github.io](https://github.com/pankajb64) 📍 [pankajb64](#) 📍 Amherst, MA

Experience

Machine Learning for Data Science Lab – Software Developer

(May 2017 to Present)

Currently involved in the Machine Learning for Data Science (MLDS) Lab, working on BirdCast, which uses CNN based image recognition to identify bird migration patterns from radar image data. We use WSRLIB, a MATLAB library for Weather Surveillance Radar written specifically for this task, coupled with a Convolutional Neural Network trained on a VGG-16 model. Contributions include scaling up the service to run on AWS clusters, and future plans include improving the performance of the CNN network. **Technologies – MATLAB, AWS S3, AWS Batch, Docker, Python.**

Play Games 24x7 Pvt. Ltd. - Software Development Engineer

(July 2013 - May 2016 – 2 years 10 months)

Worked as SDE for RummyCircle.com, one of the major products of Games24x7. Involved in Full Stack Web Development.

Accomplishments - Horizontal scaling and optimization of distributed caches using Ehcache and Terracotta, Building Restful Services for handling Cash and Reward Point transactions, strengthening application security using Spring Security, Building Event Detection System and Optimization of gameplay bot.

Technologies - Java, Spring, Spring Security, Test-NG, Jersey (JAX-RS), Hbase, Ehcache, Terracotta, Redis, Akka, MySQL, MyBatis, Node.js, JUnit.

KDE (Google Summer of Code) - Software Developer

(May-July 2011 & May – August 2012 – 7 months)

Integrated features of Facebook into KDE's micro blogging client, Choqok (Under Google Summer of Code 2012). Initially I helped built the base library LibKFBAPI, and was later involved in using the library to build features such as the ability to create a post, view posts as well as like and comment on a post.

Technologies: KDE4 based on Qt (C++ based GUI Framework). **Link** - <https://github.com/pankajb64/choqok-facebook>

Projects

Evaluating Machine Learning Techniques for Website Fingerprinting Attacks – Independent Study

(May 2017 - Present)

Currently pursuing an independent study which evaluates the assumptions made by different machine learning models for targeting Website Fingerprinting (WF) attacks on the Tor network. Current learning models include Naïve Bayes, Decision Trees and SVM, but they all employ certain restrictions (e.g., user visits only a small set of web pages, page content doesn't change over time) that aren't realistically attainable. We evaluate the assumptions made by such models and the performance in their absence. Plans include evaluating deep learning based models for such attacks, and propose adversarial examples for the same.

Technologies – Keras, Numpy, Scikit-learn, Python. **Link** - https://github.com/pankajb64/wf_attacks_evaluation

Image Captioning with Visual Attention - Machine Learning Course Project

(October – December 2016)

Incorporated an attention mechanism with gradient descent based training algorithms to build a deep learning model that learns to generate image descriptions given an unseen image, inspired by the state-of-the-art work in this field. Our results showed that attention based deep learning models can learn to produce good image captions and generate visualizations to infer how a model can focus on certain parts of the image while predicting the next word. We used a Convolutional Neural Network (CNN) to generate attention based features and a Long Short-term Memory Network (LSTM) for generating image captions.

Technologies: Keras, Theano, Python. **Link** - https://github.com/pankajb64/image_caption_using_attention

Automated Extractive Summarization of a Book using its Index - Natural Language Processing Course Project (October – December 2016)

Generating coherent summaries of long documents such as books is a relatively untouched field, with most of the research work being focused on short documents. In this Project, we looked at summarization of books which have a back-index as a novel technique for generating good summaries. We provided a new dataset for the evaluation of such summarization systems. We focused on extractive summarization using LexRank, an unsupervised learning algorithm based on eigenvector sentence centrality and compared it with the current state-of-the-art methods. The results were promising and merit future research.

Technologies: NLTK, Python. **Link** - <https://github.com/pankajb64/bookesis>

Pictionary with Jibo - HackUMass 2016

(October 2016)

Taught the robot Jibo to play Pictionary on a set of hand-drawn images. We used Jibo SDK to capture a picture and send it off to be classified by our custom-trained TensorFlow neural networks. We used openCV to detect the playing area within the board and crop off the rest of the image and make it black and white, we then passed it off to our TensorFlow model to be classified. We built the neural network using hundreds of images pulled automatically from Bing Image Search. Using TensorFlow's Image Recognition Algorithm "Inception", we successfully trained it to recognize common objects such as house, chair, etc.

Technologies: TensorFlow, Jibo SDK, OpenCV, Python. **Link** - <https://github.com/pankajb64/jibo-pictionary>

Education

University of Massachusetts Amherst

MS - Computer Science (September 2016 – PRESENT)

GPA – 3.9

Relevant Coursework – Machine Learning, Natural Language Processing, Artificial Intelligence, Systems

DA-IICT

B.Tech. - Information & Comm. Technology (2009-2013)

CPI – 8.5 (GPA – 3.3)

Relevant Coursework: Algorithms, Data Structures, Computer Networks, Databases, Operating Systems, Systems Software, Cryptography, Security Protocols, Graph Theory, Computer Organisation, Natural Computing

Skills

Interest Areas - Machine Learning, Natural Language Processing, Algorithm Design and Analysis, Software Security, Computer Networking, Information Retrieval, Artificial Intelligence

Programming Languages - Java, Python, C++, MySQL, JavaScript, Bash (Familiarity with: Scala, Ruby, C, HTML, CSS)

Development Technologies – Keras, Theano, Tensorflow, Numpy, Sklearn, Docker, AWS Batch, AWS Lambda, Spring, Android, SQLite, Hadoop, Ehcache, RabbitMQ, Redis, Akka, JDBC, MyBatis, Node.js, AngularJS, jQuery, Django, Rails, Test-NG, JUnit, Selenium, Firebase

Software Tools Used - Git, SVN, Ant, Maven, Eclipse, Android Studio, Wireshark, Vim