

# Akshay Bhatia

 [akshaybhatia10](#) |  [akshay.bhatia@knorex.com](mailto:akshay.bhatia@knorex.com) |  [akshaybhatia10.github.io](#) |  [akshay-bhatia-5283abab/](#) |

## RESEARCH INTERESTS

---

Natural Language Processing, Semi-supervised Learning, Text Generation, Text Style Transfer, Natural Language Inference

## EDUCATION

---

### Jaypee Institute of Information Technology

*B.Tech, Electronics and Communication*

Noida, India

*July. 2014 – June 2018*

**Thesis:** *Real-time control of Zigbee for Smart Shopping System using RFID*

## WORK EXPERIENCE

---

### Knorex Pte. Ltd.

*Research Scientist - NLP*

April 2019 - Present

*Pune, India*

Supervisor: Jin Yiping

- Worked on training and evaluating text classification and generation models for **Knorex KAIROS's** Brand Safety and Contextual targeting offerings. Improved overall CTR on an average advertising campaign by 12%.
- Responsible for the end-to-end development cycle of Sentinel, an Active Learning platform to build text classifiers without any labeled data. A working manuscript is available below.
- Applied Transformer models with guided decoding for the task of automatically generating high quality advertising slogans for Dynamic creative optimization, surpassing previous benchmark results by 33%.
- Supported several production ad-hoc features such as Keyword Extraction and Search Similarity for User-Defined Custom Segments for Contextual Targeting.

### CampK12

*Machine Learning Developer Intern*

May 2018 - Dec. 2018

*Gurugram, India*

Supervisor: Anshul Bhagi

- Implemented a real-time object detection and on-device scene classification model for K12 language learning Lingolens. Trained a custom YOLO model to detect ~150 objects with an mAP of 67%.
- Subsequently optimized the on-device detection and classification inference speeds by 19% with only 6% performance degradation compared to SoTA methods such as RetinaNet and SSD.
- Devised embedding and attribute-aware similarity models for the automatic question tagger system for the peer to peer question answering forum for the CampK12's [Generation Blockchain 2018 Summit](#).
- Explored the domain of Machine-assisted Assignment Grading and experimented with applying variations of neural language models and other embedding techniques to the task of Automatic Essay Scoring.

### Spikeway Technologies Pvt. Ltd.

*Machine Learning Intern*

May 2017 – July 2017

*Noida, India*

Supervisor: Praveen Kumar

- Led a 3 member team for the development and deployment of a ML-based Book Genre Classification system to classify books into their genres, based entirely on its title, without knowledge of author and origin.
- Extended trained models through containers and pods in production and directly collaborated with the backend team on the prototype of a News Article Authorship Plagiarism Checker.

## PUBLICATIONS

---

**Iterative Seed Word Selection for Weakly-Supervised Text Classification with Bayesian Error Estimation**

[Jin Yiping](#), [Akshay Bhatia](#), [Dittaya Wanvarie](#)

*EACL 2021* (Under Review)

## MANUSCRIPTS

---

**Sentinel: In-House Active Learning Platform**

[Akshay Bhatia](#), [Vishakha Kadam](#), [Jin Yiping](#), [Tho Nguyen](#)

*Paper* (In progress), 2020

## RELEVANT PROJECTS

---

### Road Network Extraction

July 2018 - Aug. 2018

Movehack Global Mobility Hackathon 2018

- A road network segmentation and extraction system using high resolution satellite imagery for reliable and low cost terrain monitoring and infrastructure quality assessment.
- Implemented and trained the U-Net architecture on the [Mnih Massachusetts](#) road dataset.
- Optimized the trained model for real-time applications with final inference speed of only 0.28 seconds on Tesla K80 GPU achieving a mask accuracy of 95% and a dice score of 65% on the validation set.

### LingoLens

Aug. 2018 - Dec. 2018

@CampK12 with Anshul Bhagi

- A multilingual language learning app for K12 students providing translations and transliterations for indoor and outdoor objects in over 20 languages.
- Implemented, trained, and deployed the scene classification model as Tensorflow Lite model supported using TFServing for on-device inference and the Object Detection model as a web API on AWS.

### Book Genre Classification

June 2017 - July 2017

@Spikeway Technologies Pvt. Ltd. with Praveen Kumar - [Live Demo](#)

- A system to predict the genre of the book given the title of a book.
- Improved upon the TFIDF and LR baselines by training a LSTM using pre-trained word2vec embeddings resulting a P/R/F1 score of 0.63/0.66/0.64.
- Deployed the trained models as a simple Flask API in a Kubernetes K8 cluster environment for a US-based client.

### Cardiac Arrhythmia ML

May 2017 - June 2017

@Udacity Machine Learning Engineer Nanodegree - Capstone Project

- A model to predict Cardiac Arrhythmia i.e. whether a patient has a “normal” or a “abnormal” heartbeat from their phonocardiogram (PCG) or heartbeat recordings.
- Implemented and trained VGG16 - a Deep Convolutional Neural Network by converting each heart sound recording(PCG) to a spectrogram image.
- Achieved an accuracy score of 66% on the [2016 PhysioNet/CinC Challenge](#) evaluation set.

### TravelCamp

Nov. 2016 - Dec. 2016

Submission for IMAD 2016 - [Live Demo](#)

- A social blog/profile web app developed for IMAD 2016.
- The web app follows RESTful approach for CRUD operations and was deployed using Heroku.
- Backend frameworks used: NodeJS, ExpressJS, MongoDB, PassportJS.

## RELEVANT COURSEWORK

---

### Certified MOOCs

- [Machine Learning Engineer Nanodegree](#) - Udacity
- [Deep Learning Foundations Nanodegree](#) - Udacity
- [Machine Learning by Stanford University](#) - Coursera
- [Introduction to Modern Application Development](#) - Hasura/IIT-Madras/NPTEL

**Other Coursework:** Object Oriented Systems and Programming, Data Analysis using R, Unix Programming, Data Structure and Algorithms

## OTHER ACHIEVEMENTS AND EXTRA-CURRICULAR ACTIVITIES

---

- Awarded the Hasura product development fellowship as one of the course toppers of [IMAD 2016](#) by Hasura/IIT-Madras.
- Selected as a Fast.ai International Fellow 2018 for Cutting Edge Deep Learning For Coders, Part 2.
- Conducted multiple meetups and Knowledge Sharing Sessions(KSS) for Research & Engineering teams across Knorex Pte. Ltd. on broad domains including [BERT](#), [Active Learning](#), [Probabilistic Programming](#).
- Co-developed, co-designed, and co-instructed CampK12's first [AI & Machine Learning](#) course over a period of 6 weeks to 15 K12 students.
- Assisted and worked with multiple clients and B2B businesses at Spikeway Technologies Pvt. Ltd. to leverage ML solutions with their products including Recommendation engines, Content Discovery, Customer Segmentation, etc.
- Worked as Internshala Student Partner; hosted and managed multiple activities and workshops across college premises to inculcate internship culture, resulting in 70 first and second-year students applying to internships through Internshala platform.

## TECHNICAL SKILLS

---

**Languages:** Java, Python, JavaScript, R

**Frameworks:** PyTorch, FastAI, OpenCV, NumPy, Scikit-learn

**Developer Tools:** Git, AWS, Docker

**Others:** PyTest, NodeJS, Flask, Latex, MongoDB

## REFERENCES

---

### **Prof. Ravinder Ahuja**

Assistant Professor, Computer Science & Engineering, Galgotias University, Noida, India(Previous: Professor, CSE, IIIT)

 [homepage](#)  [email](#)

### **Jin Yiping**

Senior Research Scientist, Knorex Pte. Ltd., Bangkok, Thailand

 [homepage](#)  [email](#)

### **Anshul Bhagi**

CEO & Co-Founder, CampK12, Gurugram, India

 [homepage](#)  [email](#)

\*\* Links in blue are hyperlinks