Bipasha Sen

Research Interests

Human Activity Forecasting, Scene Understanding, Audio-Visual Learning

Personal info

Email: bipasha.sen3195@gmail.com
Homepage: https://bipashasen.github.io/

Education

K.C.College of Engineering, University of Mumbai 2012 - 2016
B.E. in Computer Engineering (8.14/10, First Class with Distinction)

- Thesis: Reinforced and Collaborative Music Recommendation
- Relevant classes taken:
  CPC 703 Artifical Intelligence
  CPE 7023 Image Processing

Research Experience

Microsoft Research & Development 2016 - Present
Data Scientist II (Microsoft's Search and Assisted Intelligence) - Outlook Platform

- **Self-Supervised Meeting Summarization (BReSQ)**
  I am training a **self-supervised** framework called BReSQ to generate summaries of long meetings with multiple participants and speakers. **B**revity to reduce the transcript to a short latent space, **Re**levance to evaluate if the summary contains all the important points covered in the meeting, **S**pan to keep the summary from getting too short, **Q**uality to enable readability.
  I'm using a combination of Autoencoders, Generative Adversarial Networks and pretrained Question-Answering model to train BReSQ.
- **Inline Suggested Attachments**
  I'm responsible for building a **high-precision classification model** for the suggestion of relevant documents as attachments to half-composed email.
  Correct suggestion reduce the number of clicks to attach from 4 to 1.
  My tasks involve analyzing user behavior and discovering patterns on the dataset to determine the file-type, user-file-affinity of the intended attachment based on limited context (half-composed emails).
- **Meeting Insights**
  I'm responsible for building a **high-recall classification model** for the recommendation of relevant email to past and future meetings using a knowledge graph called AiGraph.
  AiGraph is generated using Outlook data consisting of meetings, emails and shared documents.
  I am building a ranker to rank a set of potential candidates for recommendation retrieved by walking AiGraph. The ranker is a classifier trained on features extracted from AiGraph.
- **Detection of Business Trips**
  Planning a trip leads to multiple reservations: Flights, Hotels etc. Keeping a track of the several bookings is a taxing job. Trips solved the problem by showing all booking relevant to a trip in a single page.
  I single-handedly developed the convoluted **algorithm** to club multiple disjoint Flight, Hotel, Bus and Cab reservation emails on Outlook to form a single logical entity representing an end-to-end trip.
- **Key Information Extraction**
  I developed a **scalable** approach for extracting key information from a long email such as Invoice amount, Account number, Due Date, without taking dependency on sender templates (airbnb.com, icici.com, etc.) by integrating Microsoft's Program Synthesis using Examples (PROSE).
  I also developed an automated pipeline to monitor the soundness of the extracted information through means of logical validation and anomaly detection.

Microsoft Research & Development December 2015
Data Scientist - Intern (Search Technology Center India)

- **Conversational Shopping Assistant Bot**
  A bot tasked for proactively engaging the users and assisting them towards placing an order.
  I developed the bot from scratch which trained itself by the means of **reinforcement learning**. I defined the optimal policy & reward and integrated Microsoft's Multi World Testing (MWT), a reinforcement learning based framework.

Publications

Reed: An Approach Towards Quickly Bootstrapping Multilingual Acoustic Models, paper
Bipasha Sen, Aditya Agarwal, Mirishkar Sai Ganesh, Anil Kumar Vuppala
Spoken Language Technology (SLT 2021)

An Approach Towards Action Recognition using Part Based Hierarchical Fusion, paper
Bipasha Sen, Aditya Agarwal
International Symposium on Visual Computing (ISVC 2020)

Microsoft Publications

Sentence Modelling for Contextual Meeting Segmentation, short-paper
Jay Paranjape, **Bipasha Sen**
Microsoft's Machine Learning and Data Sciences (MLADS 2020)

AiGraph for Meeting Insights Relevance, short-paper
Bipasha Sen, Prakash Pandey, Rajeev Gupta, Vipin Vangala
Microsoft's Machine Learning and Data Sciences (MLADS 2020)

Major projects

Reinforced and Collaborative Music Recommendation 2016
Undergraduate Thesis
Developed an agent that recommends music from the song-library on the phone. The agent continuously learns and evolves based on collaborative (users wiht similar behavioral pattern) feedback.

Anterior Segment Imaging (MIT Media Lab's REDX Camp) 2015
REDX is an interdisciplinary platform to enable collaboration between world-renowned medical professionals and engineers to build solutions for society's most pressing eye related healthcare challenges. Developed a low-cost, solid-state device with no moving parts, as a replacement for heavy and bulky Ophthalmic Slit Lamp, to capture and reconstruct 3D model of the cornea (anterior segment of the eye) for differentiating between a healthy and unhealthy cornea.

TheBhaad: Cloud-Based Group-Oriented file sharing network (video) 2014
Single-handedly developed a fully-fledged cloud-based file sharing network with windows like user-interface. Features: Search, Contacts, Groups (Classrooms), Personalized Document Alignment.

Awards and Honors

Invited for talk at MLADS on Quick Bootstrapping of Multilingual Models	July 2020
3rd in Microsoft One Week Hackathon - Mobile Endpoint (3k+ participants)	August 2016
126th in TCS CodeVita '15 Round 2 (19800+ participants)	February 2016
Best Student of the Year (Out of 600+ graduating students)	February 2016
Best Entrepreneur (For founding TheBhaad and hosting 5000+ users)	March 2015

Skills

Languages	Python, Spark.net, C#, C/C++, HTML, CSS, jQuery
Framework	Pytorch, Tensorflow, scikit-learn

Extra-Curricular

I am a musician: vocalist, guitarist and composer. I've toured around India along with my previous band, Andrometa. I've also travelled to 6 countries - 11 states solo and met 70+ bands (180+ artists).