

Shramay Palta

UNIVERSITY OF MARYLAND, COLLEGE PARK

4108, Brendan Iribe Center for Computer Science and Engineering

+1 (443) 858-4922 | ✉ spalta@umd.edu | 🏠 shramay-palta.github.io | 🌐 shramay-palta | in shramay-palta | 🐦 @PaltaShramay | 📧 Shramay Palta

Education

University of Maryland College Park

DOCTOR OF PHILOSOPHY, COMPUTER SCIENCE. **GPA: 4.00/4.00**

Working with **Professor Rachel Rudinger**.

College Park, Maryland, USA

AUGUST 2021-PRESENT

Birla Institute of Technology and Science, Pilani (BITS Pilani)

BACHELOR OF ENGINEERING, ELECTRICAL AND ELECTRONICS ENGINEERING . **FIRST DIVISION**

Thesis Supervisor: **Dr. Ashok Agrawala (UMD)** and **Dr. Navneet Gupta (BITS)**.

Pilani, India

AUGUST 2017-MAY 2021

Research and Work Experience

Computational Linguistics and Information Processing (CLIP) Lab

GRADUATE STUDENT UNDER PROF. RACHEL RUDINGER

NOVEMBER 2021 - PRESENT

- Investigating modern-day **commonsense reasoning** NLP Models and datasets like **Delphi**, **COMET**, **ATOMIC**, and **CommonsenseQA** to determine if they have an implicit or explicit **cultural bias** baked into them.
- Preparing a **dataset** of test bed questions that can be used to stress test these models using examples of cultural and social norms, material and physical differences to prove such biases.
- Testing multiple models like **BERT-base**, **BERT-Large**, **RoBERTa** to determine the affect on performance by measuring how the answer varies when the cultural context is explicitly or implicitly specified in the question and when it is not.

College Park, Maryland, USA

Human-Data Interaction Group, University of Maryland

GRADUATE RESEARCH ASSISTANT UNDER PROF. LEO ZHICHENG LIU

SEPTEMBER 2021 - MARCH 2022

- Using Natural Language Processing Techniques to harvest design feedback from visualization comments on social media platforms like Reddit.
- Crawled **~26400** posts from **/r/dataisbeautiful** to get **~2.5 Million** comments to be used for the study.
- A qualitative study of **600** sampled comments was used to derive meaningful taxonomy for prominent dimensions and categories for online visualization feedback.
- Used three **traditional classifiers**: **Naïve Bayes**, **Decision Tree**, and **Random Forest**, and a base **BERT** model for classification.

College Park, Maryland, USA

Maryland Information and Network Dynamics (MIND) Lab, University of Maryland

RESEARCH ASSISTANT UNDER PROF. ASHOK AGRAWALA (UNDERGRADUATE RESEARCH THESIS)

MAY 2020 - FEBRUARY 2021

- Analyzing the spread of **COVID-19 and Flu virus** on campus using location and breathing data collected from **Spire Tags**. Part of the **PROMETHEUS** Project in collaboration with the School of Public Health.
- Designed algorithms to efficiently determine **proximity pairs for contact tracing** on a data set of more than **500K** data points. Generated proximity pairs between two users within a distance of 3 meters and within a span of 15 minutes from each other.
- Developed algorithms for **time series analysis** for **segmentation of breathing data** to identify singular breathing events and anomalies like talking, sneezing, laughter, cough etc.

College Park, Maryland, USA

Global Health Centre, Graduate Institute of International and Development Studies

RESEARCH INTERN UNDER DR. AMANDEEP GILL, EXECUTIVE DIRECTOR, UNSG'S PANEL ON DIGITAL COOPERATION

MAY 2020 - OCTOBER 2020

- The International Digital Health and AI Research Collaborative (I-DAIR) Project seeks to advance the **UN Secretary General's (UNSG)** High-level Panel on Digital Cooperation's recommendations related to digital health, and targets set at the **World Health Organization (WHO)** on universal and quality health coverage.
- Researched the role of **micro-narratives** as proxy variables to fill in missing data and used **natural language techniques** to study the social, health, and impacts of the **COVID-19** crisis on various sections of the society.

Geneva, Switzerland

TurnoutNow LLC

DATA SCIENCE INTERN

MAY 2019-JULY 2019

- Analyzed how Big Data from **25000+ IoT BLE Beacons** is created and captured by over **900+ IoT data capture devices (Session App)**.
- Used **vert.x-core framework** for capturing and storing this **Big Data** into **Redis**. Used Java to process it in real-time and **MongoDB** for archiving.
- Used **R** to generate **real-time recommendations** based on the attendee's live location on the show floor.
- The data was then analyzed in **Power BI Dashboards** which displayed an overall summary of the event with **natural language insights**.
- Used real-time natural language generation tools with live data connections and generated narratives for end users. These narratives provided controlled insights based on user inputs in a concise natural language text.

Lancaster, Pennsylvania, USA

Publications

Investigating Information Inconsistency in Multilingual Open-Domain Question Answering

SHRAMAY PALTA, HAOZHE AN, YIFAN YANG, SHUAIYI HUANG, MAHARSHI GOR

Activities

Reviewer: EMNLP 2022: Sentiment Analysis, Stylistic Analysis, and Argument Mining track

Projects

Analyzing Inconsistencies in Multilingual Open-Domain QA

College Park, Maryland

PROF. JORDAN BOYD-GRABER, DEPT. OF COMPUTER SCIENCE, UNIVERSITY OF MARYLAND

MARCH 2022- PRESENT

- Investigating whether **multilingual question answering** can potentially expose users to unreliable information through **cultural differences, divergent national laws, or uneven resources**.
- Analyzing if **different retriever models** present different passages—and answers—given the same question in different languages on **TyDi QA** and **XOR-TyDi QA**, two typologically diverse multilingual QA datasets
- Different answers reveal valuable information about per-language **resources disparity, and linguistic variation**.

Detecting Dietary Activity with Wearable Earphones

College Park, Maryland

PROF. NIRUPAM ROY, DEPT. OF COMPUTER SCIENCE, UNIVERSITY OF MARYLAND

SEPTEMBER 2021- DECEMBER 2021

- Developed a technique to detect the types of food being consumed using **eSense**, a consumer wireless earphone device by Bell Labs.
- Used data from the microphone to **detect chewing activity** and classify the food being chewed as **solid, liquid or semi-solid**.
- Choosing **Convolutional Neural Networks (CNNs)** as the feature extractor, used standard **Conv2d** layers and a standard training and evaluation procedure to train the models both with and without the **Mel Spectrogram**.

Economic Psychology: Stock Market Prediction using BERT

BITS Pilani, India

DR. RAJNEESH CHOUBISA, DEPT. OF HSS, BITS PILANI

JANUARY 2021- MAY 2021

- Implemented a **neural network** for utilizing information in **SEC 8-K forms** for predicting the movement of the **S&P 500 index**.
- Used **BERT** for capturing the contextual information in the form of two methods: **Masked Language Modeling (MLM)** and **Next Sentence Predicting (NSP)**.

Darknet Insights using R and Python

BITS Pilani, India

DR. VISHAL GUPTA, DEPT. OF CSIS, BITS PILANI

AUGUST 2019-DECEMBER 2019

- Used information from DNS queries to predict a **DDoS attack** from Darknet data from **Center for Applied Internet Data Analysis (CAIDA)** supercomputer servers of the **University of California, San Diego (UCSD)**.
- Implemented **Python scripts for feature extraction** (like TTL, IP length, Packet Count etc.) and **CAIDA's internal tool, corsaro** for large scale analysis of trace data.
- Used **vector quantization algorithms including K-means and EM** on the extracted features to predict DDoS attacks.

Skills

Languages	Python, R, SQL, Linux/Unix shell, Java, C++, C, Assembly Language.
Tools	Pandas, NumPy, NLTK, spacy, Keras, TensorFlow, PyTorch, scikit-learn, Matplotlib, Jupyter, Git, \LaTeX , MATLAB, MySQL.
Key Courses	CommonSense Reasoning and Natural Language Understanding, Natural Language Processing, How and Why AI Answers Questions, Human AI Interaction, Advanced Numerical Optimization.
OS	MacOS, Linux, Windows, FreeBSD.

Achievements

2022	Dean's Fellowship Award: Awarded the Graduate School Dean's Fellowship Award for outstanding academic achievement.
2021	Dean's Fellowship and Chair's Fellowship Award: Awarded the Graduate School Dean's Fellowship and the Chair's Fellowship Award for outstanding academic achievement.
2015	National Merit Holder: One of the top 0.1% scorers across India in the board examinations conducted by Central Board of Secondary Education; Received Letter of Honour from the HRD Minister, Govt. of India.
2010	World Robot Olympiad: Represented India at the World Robot Olympiad held in Manila, Philippines and managed to secure a world rank of 31 in my category.
2010	Indian Robot Olympiad: Awarded the 1st Runners Up Award in my category for the north chapter.