

# Shubhashis Roy Dipta

No visa sponsorship required; Green Card holder.



[sroydip1@umbc.edu](mailto:sroydip1@umbc.edu) | [in/dipta007](https://in.dipta007) | [RoyDipta.com](https://RoyDipta.com) | [dipta007](https://github.com/dipta007) | [\(443\) 889-3961](tel:(443)889-3961)

## EDUCATION

**Ph.D. in Computer Science**, UNIVERSITY OF MARYLAND, BALTIMORE COUNTY

Jan 2021 – December 2026 (Expected)

Specialization: **Decomposition-based Multimodal Reasoning, Tool Calling LLM, Vision Language Model (VLM), Natural Language Processing (NLP)**

- Phi Kappa Phi Award - **Top 10% of STEM**. GPA 4.00/4.00 – **Top 1% of class**.

## EXPERIENCE

**Applied Research Scientist Intern**, AMAZON

June 2025 – Aug 2025

Tech Stack: PyTorch, Agentic LLM, Multi-Stage SFT, PPO, GRPO, RL, Weight & Biases

- Developed multi-stage continual training to inject rules, outperforming  $\tau$ -Bench baselines by **20 points** without in-context rules while decreasing the **context length by 40%**.
- Introduced novel **PolicyRecall** reward with Jaccard score and hallucination penalty in GRPO for precise policy adherence in LLM.
- Developed synthetic benchmark,  **$\tau$ -Override**, to evaluate the ability of LLMs to override learned rules.

**Machine Learning Research Intern**, SCALE.AI

June 2024 – Aug 2024

Tech Stack: PyTorch, DeepSpeed, Hugging Face, Kubernetes, dagster, RLHF, SQL, Weight & Biases

- **RLHF text2sql**: Used Online KTO (a novel method) and data augmentation to improve the **BIRD** benchmark accuracy by **8 points** over the same size SFT model (Paper In ARR review).
- **Auto Eval**: Implemented AutoCOT, Self Critique to improve the current Auto Eval system by 6 points (**F1 score: 83 → 89**).
- **Many-Shot text2sql**: Implemented Many-Shot *text2sql* system, improving the total accuracy by **8%** and per use case accuracy by **12%**.

**Graduate Research Assistant**, UNIVERSITY OF MARYLAND, BALTIMORE COUNTY

Jan 2021 – Present

Tech Stack: PyTorch, Python, Hugging Face, Scikit-Learn, LLM, SQL, Spacy, NLTK, Pandas, Numpy, Matplotlib, Seaborn, Weight & Biases, Hydra

- **Video-Text Retrieval**: Enhancing Zero-Shot Multilingual Video-Text Retrieval through Outcome and Intent-Driven Modeling Utilizing Only Visual Features.
- **Event Modeling**: Achieved **8.5% improvement** over prior state-of-the-art approaches in **2 datasets** and **across 4 evaluation metrics** by designing a novel, hierarchical, semi-supervised event modeling framework. (Oral Presentation on [\\*SEM 23](#), [ACL](#))
- **Multimodal Counterfactual**: Pioneering **first-of-its-kind** multimodal counterfactual dataset (**8k+ real-life events**), merging text and images for nuanced alternate timeline, a novel contribution to counterfactual reasoning and multimodal real-life event understanding.

**Machine Learning Engineer**, BACKPACKBANG.COM

Oct 2018 – Mar 2019

Tech Stack: PyTorch, Keras, Hugging Face, Large Language Model, NLTK, Python, MySQL, Elasticsearch, AWS EC2, Node.js, React

- **Boosted the sale by  $\approx 23\%$**  by improving existing product recommendation system using *Product2Vec* embedding.
- **Decreased server cost by  $\approx 10\%$**  by implementing an AWS Lambda-based ML pipeline for online learning.

**Founder & Chief Technology Officer**, UNISHOPR.COM

Jan 2019 – Jan 2021

Tech Stack: Python, Node.js, React, PostgreSQL, GraphQL, AWS Lambda, AWS Lightsail

- Single-handedly led a cross-functional team of 10, achieving **1,000+** active e-commerce users and **\$100,000+/month** in orders.

## TOP PUBLICATIONS

For a comprehensive list, visit my [Google Scholar](#) Profile

1. [Shubhashis Roy Dipta](#), and Francis Ferraro. "If We May De-Presuppose: Robustly Verifying Claims through Presupposition-Free Question Decomposition" Proceedings of the 14th Joint Conference on Lexical and Computational Semantics, EMNLP (2025)
2. [Shubhashis Roy Dipta](#) "VC-Inspector: Advancing Reference-free Evaluation of Video Captions with Factual Analysis" (On Review)
3. [Shubhashis Roy Dipta](#), and Francis Ferraro. "Q2E: Query-to-Event Decomposition for Zero-Shot Multilingual Text-to-Video Retrieval." (On Review)
4. [Shubhashis Roy Dipta](#), Mehdi Rezaee, and Francis Ferraro. "Semantically-informed Hierarchical Event Modeling." Proceedings of the 12th Joint Conference on Lexical and Computational Semantics, ACL (2023)
5. [Shubhashis Roy Dipta](#), and Sai Vallurupalli. "UMBCLU at SemEval-2024 Task 1A and 1C: Semantic Textual Relatedness with and without machine translation." Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2024), NAACL (2024)
6. [Shubhashis Roy Dipta](#), and Sadat Shahriar. "HU at SemEval-2024 Task 8A: Can Contrastive Learning Learn Embeddings to Detect Machine-Generated Text?." Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2024), NAACL (2024)

## RELEVANT PROJECTS

For a comprehensive list, visit my [Portfolio](#)

**Bird Chirping Identification** [↗](#)

Tech Stack: PyTorch, Keras, TensorFlow, Python, Matplotlib, Seaborn, EfficientNet

- Secured **Bronze Medal** in Kaggle (**Top 6%**) by using EfficientNet on the spectrogram with a weighted ensemble of framed timespan.

**Amazon [Crawler] [Search Engine]**

Tech Stack: Elasticsearch, Node.js, Express.js, Python, Multithreading, GCP

- Designed a distributed data pipeline using GCE to extract and index over **10M products**, enhancing resource-efficient data retrieval.

## PROFESSIONAL SERVICES

- **Competitions**: 2 International Robotics Competitions (URC, USA; ERC, Poland), 2 ACM-ICPC, 25+ National Programming Competitions.