

Vedant Palit

(+91) 9163389534 | vedantpalit@kgpian.iitkgp.ac.in | [linkedin.com/in/vedant-palit](https://www.linkedin.com/in/vedant-palit) | github.com/vedantpalit | [google scholar](#)

EDUCATION

| | |
|--|-------------|
| Indian Institute of Technology(IIT) Kharagpur, India <i>Btech in Industrial & Systems Engineering and Mtech in Financial Engineering</i> | 2021 – 2026 |
| Birla High School <i>Subjects: English, Physics, Chemistry, Mathematics, Computer Science</i> | 2008 – 2021 |

RESEARCH EXPERIENCE

| | |
|--|---|
| Mechanistics Interpretability of Vision-Language Models <i>Stanford University</i> | Stanford, CA (Remote) <i>Dec 2023 – Present</i> |
| <ul style="list-style-type: none">Working on expanding arXiv:2308.14179 for extensive analysis of neural activations inside VL Architectures.Involved in reproducing the claims of arXiv:2310.18969 alongside additional ablation studies and adaptations to Vision-Language Models. | |
| Causal Intervention on the BLIP Architecture <i>Carnegie Mellon University</i> | Pittsburgh, PA (Remote) <i>April 2023 – Sep 2023</i> |
| <ul style="list-style-type: none">Created a pipeline adapting causal mediation analysis to interpret blackbox architectures of VL transformers.Implemented the method on the BLIP transformer and used the COCO-QA dataset to study the effect of various layers on the final outputs. | |
| Wellness Dimensions Benchmark for Explainability of LMs <i>University of Maryland, Baltimore</i> | Baltimore, MD (Remote) <i>Nov 2022 – Dec 2023</i> |
| <ul style="list-style-type: none">Trained various general and domain-specific models for suicide risk assessment, using the gamblers and cross entropy loss functions on annotated datasets containing social media posts classified into 6 different wellness dimensions.Utilised singular value decomposition to analyse the impact of the loss function on the attention scores of the models. | |
| Knowledge Graph Guided Semantic Evaluation of LMs For User Trust <i>University of South Carolina</i> | Columbia, SC (Remote) <i>Feb 2023 – March 2023</i> |
| <ul style="list-style-type: none">Developed a novel evaluation method to measure error in reconstruction of masked knowledge graph structures from outputs by LLMs.Analysed and benchmarked the performance of GPT-3.5, GPT-J and GPT-NeoX in reconstructing KG paths using the evaluation metric. | |

PUBLICATIONS

| | |
|--|--|
| Towards Vision-Language Mechanistic Interpretability: A Causal Tracing Tool for BLIP [Link] <i>Published at the ICCV 2023 Workshop on Closing The Loop Between Vision and Language - First Author</i> | |
| Knowledge Graph Guided Semantic Evaluation of Language Models For User Trust [Link] <i>Published and Presented at the IEEE Conference on Artificial Intelligence 2023</i> | |
| Benchmark of Wellness Dimensions for Robustness and Explainability Evaluation in LMs <i>Under Review at the ACL CLPsych Workshop 2024</i> | |

PROJECTS

| | |
|--|--|
| Ocean-Freight Optimization [Report] <i>Semester Project for Operations Research Laboratory</i> | Kharagpur, India <i>Jan 2023 – April 2023</i> |
| <ul style="list-style-type: none">Utilised and adapted the concept of capacity constrained vehicle routing to determine the optimal path to be taken by ships from a shipping company, using GurobiPy and CPLEX. | |
| Knowledge Graph Generation from Scraped Text [Repo] <i>Task Round for the AIISC Internship</i> | Columbia, SC (Remote) <i>August 2022 – September 2022</i> |
| <ul style="list-style-type: none">Created an automated text extraction system, for custom information extraction from the user's desired wikipedia page followed by entity generation through parsing, named entity correlation and KG generation. | |

Image Stitching and RANSAC Line Fitting [Repo]

Kharagpur, India

Task Round for the Autonomous Ground Vehicles Group, IIT Kharagpur

May 2022 – July 2022

- Constructed a pipeline of the image stitching through brute-force matching as well as KNN Association of keypoints in two images of the same scene taken from close yet different angles.

Lane Detection [Repo]

Kharagpur, India

Task Round for the Autonomous Ground Vehicles Group, IIT Kharagpur

May 2022 – July 2022

- Designed an algorithmic implementation of lane detection utilising a Gaussian blur and Gamma Correction for brightness adjustment, followed by canny edge and contour detection to detect roads in real-life images.

Kalman Filter Trajectory Determination [Repo]

Kharagpur, India

Task Round for the Autonomous Ground Vehicles Group, IIT Kharagpur

May 2022 – July 2022

- Developed a trajectory determination program based on the kalman filter algorithm utilising information from acceleration and position sensors to predict the next position of a vehicle.

RELEVANT COUREWORK

- **University:** Machine Learning: Foundations and Applications(AI42001), Operations Research-II(IM21204), Operations Research Laboratory(IM29204), Transform Calculus(MA20202), Probability and Statistics(MA20205), Operations Research-I(IM21201), Programming and Data Structures(CS10003), Programming and Data Structures Laboratory(CS19003), Linear Algebra, Numerical and Complex Analysis(MA11004), Advanced Calculus(MA11003)
- **MOOCs:** Natural Language Processing with Deep Learning(CS224N), Introduction to Algorithms(MIT 6.006), Design and Analysis of Algorithms(MIT 6.046J), Neural Networks and Deep Learning, Machine Learning Specialization, Machine Learning

POSITIONS OF RESPONSIBILITY

Subhead — Associate [Drive]

Kharagpur, India

Business Club, IIT Kharagpur

Dec 2021 – Dec 2022

- Conducted a special Whitepaper Presentation on the theory of Recommender Systems and Natural Language Processing, under the guidance of the Analytics Wing of the Club.
- Organized the introductory knowledge session on Business Strategy, imparting essential insights and key concepts to a batch of 36 associates.

TECHNICAL SKILLS

Programming Languages: C/C++, Python, MATLAB

ML-DL: TensorFlow, Pytorch, Torchvision, Sklearn, Caffe

CV-NLP: Transformers, OpenCV, PIL, NLTK, Spacy

Miscellaneous: Mysql, LaTeX, HTML, Markdown, Git

AWARDS AND ACHIEVEMENTS

JEE Advanced: Secured an AIR of 4472 out of 150k candidates appearing in Joint Entrance Examination Advanced, 2021.

JEE Mains: Secured an AIR of 8241 out of 930k candidates appearing in Joint Entrance Examination Mains, 2021.

WBJEE: Secured a state rank of 99 out of 70k candidates appearing in West Bengal Joint Entrance Examination, 2021.

Technical Writing: Writer of a series of blogs reviewing papers on ML, DL and AI. [Medium]

NSS Volunteer: Recipient of the gold medal for exceptional service work as an active participant in cleanliness drives, clothes distribution drives and education camps conducted by the NSS in villages near Kharagpur.

Scientific Forum: Selected as a delegate of India at the Asia Pacific Forum for Science Talented 2019.

Case Study: Stood 1st amongst 5000+ participants in the BITS APOGEE, CaseQuesta challenge 2022.