INDERJEET NAIR

📞 +1-734-489-4505 | ✉ inderjeetnair1@gmail.com | ⌂ inderjeetnair | ⌂ inderjeetnair.github.io

INTERESTS

Natural Language Processing, Information Extraction, Document Intelligence, Machine Learning

EDUCATION

- **Pursuing Ph.D. in Computer Science, University of Michigan, Ann Arbor**   Aug 2023 - Current
  Advisor: Prof. Lu Wang

- **Bachelor of Technology, Indian Institute of Technology Bombay**   July 2017 - April 2021
  Major in Electrical Engineering and Minor in Computer Science and Engineering
  Cumulative GPA: **9.56/10.00**

PUBLICATIONS

[1] **Inderjeet Nair**, Jiaye Tan, Xiaotian Su, Anne Gere, Xu Wang, Lu Wang, **Learning to Generate Writing Feedback via Language Model Simulated Student Revisions**, *Under submission.*

[2] **Inderjeet Nair**, Lu Wang, **MIDGARD: Self-Consistency Using Minimum Description Length for Structured Commonsense Reasoning**, In *Proceedings of the Association for Computational Linguistics* (ACL '24), 2024. [LINK]

[3] **Inderjeet Nair***, Shwetha S.*, Apoorv Saxena, Koustava Goswami, **Drilling Down into the Discourse Structure with LLMs for Long Document Question Answering**, In *Findings of the Empirical Methods in Natural Language Processing* (EMNLP '23), 2023. [LINK]

[4] **Inderjeet Nair**, Natwar Modani. **Exploiting Language Characteristics for Legal Domain-Specific Language Model Pretraining**. In *Findings of the European Chapter of the Association For Computational Linguistics* (EACL '23), 2023. [LINK]

[5] **Inderjeet Nair**, Aparna Garimella, Balaji Vasan Srinivasan, Natwar Modani, Niyati Chhaya, Srikrishna Karanam, Sumit Shekhar. **A Neural CRF-based Hierarchical Approach for Linear Text Segmentation**. In *Findings of the European Chapter of the Association For Computational Linguistics* (EACL '23), 2023. [LINK]

[6] Aishwarya Agarwal*, Anuj Srivastava*, **Inderjeet Nair***, Swasti Shreya Mishra*, Vineeth Dorna, Sharmila Reddy Nangi, Balaji Vasan Srinivasan. **SketchBuddy: Context-Aware Sketch Enrichment and Enhancement.** In *Proceedings of the 14th Conference on ACM Multimedia Systems* (MMSys '23), 2023. [LINK]

[7] Shubham Anand Jain*, Rohan Shah*, Sanit Gupta[†], Denil Mehta[†], **Inderjeet Nair**[†], Jian Vora[†], Sushil Khyalia, Sourav Das, Vinay J Ribeiro, Shivaram Kalyanakrishnan. **PAC Mode Estimation using PPR Martingale Confidence Sequences.** In *International Conference on Artificial Intelligence and Statistics* (AISTATS '22), 2022. [LINK]

[8] Natwar Modani, Anurag Maurya, Gaurav Verma, **Inderjeet Nair**, Vaidehi Patil, Anirudh Kanfade. **Detecting Document Versions and Their Ordering in a Collection**. In *International Conference on Web Information Systems Engineering 2021* (WISE '21), pages 405-419, Springer, Cham, 2021. [LINK] [**Best Paper Runner-up Award**]

PATENTS

[1] **Inderjeet Nair**, Anirudh Phukan, Aravind Veluri, Lakshya J., Mohar Kundu, Akhash Amarnath, Niyati Chhaya, Sumit Shekhar. **Reflowing Infographics for Enhanced Cross-Device Consumption** [Accepted For Filing]

[2] **Inderjeet Nair**, Akshay Singhal, Kumud Lakara, Pritika Ramu, Vikas Balani, Anandhavelu N. **Minimally Guided Semantic Extraction** [Accepted for Filing]

[3] **Inderjeet Nair**, Natwar Modani. **Exploiting Legal Domain Characteristics for Legal Language Model Pretraining** [Accepted for Filing]

[4] **Inderjeet Nair**, Natwar Modani. **Integrated Reading Experience for Contracts and their Amendments** [Filed] (US Patent App. 17/954,558)

[5] Ayush Maheshwari, **Inderjeet Nair**, Navita Goyal, Natwar Modani, Ani Nenkova. **Assisted Review of Text Content using a Machine Learning Model** [Filed] (US Patent App. 17/549,270)

[6] Natwar Modani, Vaidehi Patil, **Inderjeet Nair**, Gaurav Verma, Anurag Maurya, Anirudh Kanfade. **Systems for Generating Indications of Relationships between Electronic Documents** [Filed] (US Patent App. 17/534,744)

WORK EXPERIENCE

- **Research Associate - Adobe Inc. (Big Data Intelligence Lab)**   Jul 2021 - Aug 2023
  Group: Multimodal Content Group   *Bangalore, India*
  - ∗ Engaged in several projects involving Document Intelligence, Natural Language Processing, Legal AI
  - ∗ Ideated several industrial use cases that were well received by business executives and developed accurate and efficient machine learning algorithms for them
  - ∗ Submitted papers, filed patents, gave talks on my projects, and ran tutorials that everyone found helpful

- **Research Intern - Adobe Inc. (Big Data Intelligence Lab)**   Apr 2020 - Jul 2020
  Topic: Document Families: Finding Lexical and Semantic Relations between Documents   *Natwar Modani, Gaurav Verma*
  - ∗ Designed a robust and performant framework for detecting version, prerequisite, and similarity relations among documents
  - ∗ Formulated a **topic-modeling** based approach to determine dependency relations between documents without using an external knowledge base
  - ∗ Bagged **Best Paper Runner-up Award at WISE '21** and filed a patent

- **Production Engineer - StampMyVisa**   May 2019 - Aug 2019
  - ∗ Contributed extensively to the development of a Cross-platform compatible mobile application with integration of **Firebase messaging service** and **Razorpay** in **React-Native**
  - ∗ Leveraged **Nodejs**, **MongoDB** and **Express** in the development of the backend server and implemented the Web application portal in **React-Typescript**

KEY PROJECTS

- **A Neural CRF-based Hierarchical Approach for Linear Text Segmentation**   May 2022 - August 2022
  Adobe Research
  - ∗ Formulated a novel approach for inducing **linear segmentation** based on **hierarchical topical segmentation**
  - ∗ Proposed first supervised technique for hierarchical segmentation using **CRF** to explicitly model the statistical dependency between a node and its constituent child nodes
  - ∗ Utilized **Chomksy Normal Form (CNF)** theory to design an algorithm to convert a generic hierarchical structure to **Binarized form** and vice-versa to ensure tractable computation of tree structure CRF objective
  - ∗ Compiled a vast corpus of over **700K** Wikipedia articles using automated methods for providing ground truth hierarchical structures and designed a novel data augmentation technique to boost model performance

- **Legal Domain-Specific Language Model Pretraining**   May 2022 - August 2022
  Adobe Research
  - ∗ Proposed a novel approach involving **legal domain-specific objectives** for language model pretraining unlike domain agnostic approaches like MLM and Auto-Regression over legal corpus
  - ∗ Formulated an objective to optimize the model in understanding the components of the **templatized language** used in legal frameworks and proposed a data curation strategy for this task
  - ∗ Demonstrated superiority of this approach over several standard legal downstream tasks

- **Towards Improved Document Consumption**   August 2021 - Present
  Adobe Research
  - ∗ Envisioned a novel system that personalizes a document based on user specifications, involves innovative navigation technologies, and allows the user to adaptively personalize it with minimal user interactions
  - ∗ Bagged **special-mention award** at Adobe's Hackweek competition which was participated by over 70 teams
  - ∗ Well received by several Business Unit leaders at Adobe and plans are underway for **productization**

- **Assistance in Contract Review**   August 2021 - Present
  Adobe Research
  - ∗ Formulated machine learning systems for several use cases pertinent to contract review:
    - − Integrated reading experience for master contract and its amendments
    - − Automated review of agreements by detecting rights and responsibilities and risky clauses
    - − Developed legal language model that boosts performance across several legal downstream tasks
  - ∗ Filed **3 patents** and submitted a paper to a top-tier conference

- **Estimating the Winner of Elections using Confidence Sequences**   Jan 2021 - July 2021
  Prof. Shivaram Kalyankrishnan   *Dept. of CSE, IIT Bombay*
    * Formulated mode estimation algorithm using **Prior-Posterior Ratio martingale confidence sequences** to validate the winner of an election in the fewest number of samples
    * Validated the effectiveness of our approach by performing extensive empirical analysis in the task of winner estimation of an election in single and multi-constituency setting
    * Derived tight asymptotic bounds for our sample complexity and showed that our stopping rule is asymptotically optimal

- **Generative Modelling for Joint Task of Classification and Anomaly Setection**   August 2020 - July 2021
  Prof. Suyash Awate   *Dept. of CSE, IIT Bombay*
    * Formulated a **deep learning generative mixture model** which employs **min-max GAN learning** along with **Expectation Maximization** algorithm to model the distribution associated with the input data
    * Incorporated innovative techniques like **Noise Stabilization** and **Spectral Normalization** to mitigate the problem of discriminator overfitting when the number of samples was limited
    * Demonstrated the effectiveness of our approach across several real-world datasets like MVTEC, BCCD, etc.

TEACHING EXPERIENCE

- Undergraduate Teaching Assistant for *Differential Equations (MA207)* course at IIT Bombay   2019
- Undergraduate Teaching Assistant for *Physics of Electricity and Magnetism (PH107)* course at IIT Bombay   2019

SKILLS

- **Languages**   Python, C++, Javascript, Typescript, MATLAB, VHDL
- **Packages and Frameworks**   Pytorch, Pytorch Lightning, Huggingface Transformers and Datasets, Tensorflow, ReactJS, React Native, NodeJS, Express, MongoDB, Redux

SCHOLASTIC ACHIEVEMENTS

- Secured 5th rank in MHT-CET among 3,89,520 participants   2017

- Offered fellowship under Kishore Vaigyanik Protsahan Yojana (KVYP), conducted by Department of Science and Technology, Government of India, with **All India Rank 249**   2017

- Secured **99.195** Percentile in JEE Advanced 2017 out of 1,59,540 candidates and placed in top **0.1%** of JEE Mains 2017 among 12,00,000 applicants   2017

- Ranked in **national top 1%** in National Standard Examination in Physics and placed in **statewise top 1%** in National Standard Examination in Chemistry   2016-17