# Noveen Sachdeva

✉ nosachde@ucsd.edu | 🏠 www.noveens.com | ⭘ noveens | in noveensachdeva | Google Scholar

## Education

**UC San Diego**                                                                                                               *USA*
Ph.D. in Computer Science                                                                               *2020 - PRESENT*
*Advisor:* Prof. Julian McAuley , *GPA:* 4.0

**IIIT Hyderabad**                                                                                                           *India*
B.Tech (Hons.) & M.S in Computer Science                                                                *2015 - 2020*
*GPA:* 9.75/10 (M.S)

## Experience

**Pinterest Labs**                                                                                           *San Francisco, CA*
Research Intern                                                                                         *Jun. 2021 - PRESENT*
- Working towards improving the ads quality and ads ranking systems at Pinterest.
- Project involves designing smart negative sampling strategies for CTR prediction models.

**Microsoft Research**                                                                                       *Bangalore, India*
Research Intern                                                                                     *Jan. 2020 - Jun. 2020*
- Worked with Dr. Manik Varma and his group on building better machine learning algorithms at the million-scale (Extreme Classification).
- Formulated a scalable, GCN-inspired algorithm which exploits label-label correlation patterns to massively improve tail-label performance.
- Full paper accepted for publication at WWW '21.

**UC San Diego**                                                                                               *San Diego, CA*
Research Assistant                                                                                   *Aug. 2019 - Nov. 2019*
- Worked with Prof. Julian McAuley on aspects of applied machine learning, specifically in the context of NLP and recommender systems.
- Ascertained a highly relevant problem in existing recommender systems that exploit textual reviews for rating prediction, and generalize it.
- Wrote a paper about the realized problem and possible fixes under different scenarios. Paper accepted for publication at SIGIR '20.

**Cornell University**                                                                                            *Ithaca, NY*
Research Assistant                                                                                    *Jun. 2019 - Jul. 2019*
- Worked with Prof. Thorsten Joachims and his group at the intersection of causal inference, counterfactual learning, and reinforcement learning.
- Contributed to a $1 Million project (NSF #1513692) on making off-policy learning from biased, logged contextual-bandit data more robust.
- Formalized a highly relevant problem and generalized different estimators. Paper accepted for publication at KDD '20 (Research Track).

**PwC - PricewaterhouseCoopers**                                                                                  *Tampa, FL*
Data Science Intern (Remote)                                                                          *Aug. 2018 - Nov. 2018*
- Worked with the data science and innovation team on clause extraction from sensitive legal documents for top clients in the US.
- Formulated a de-generate pipeline and compared different statistical and deep-learning based models for the given task.
- Reduced task time from days to a few hours which enabled PwC to get new clients in the legal sector.

**National Research Council of Italy**                                                                          *Rende, Italy*
Research Assistant                                                                                    *May. 2018 - Jul. 2018*
- Worked with senior researcher, Dr. Giuseppe Manco on building novel and better systems suited for the task of next-item recommendation.
- Devised a taxonomy of VAE models for collaborative filtering, demonstrating huge gains over existing state-of-the-art on real-world datasets.
- The project's findings were later published at top data-mining conference, WSDM '19.

**Google Summer of Code**                                                                                 *Nüremberg, Germany*
ownCloud                                                                                             *May. 2017 - Aug. 2017*
- Implemented a JS-library, with exhaustive unit-tests and swagger-documentation, which works both on Node.JS and browser.
- Presented talk at annual ownCloud conference at Nüremberg, Germany. Project led to real-world gains in usage of ownCloud.

**IIIT Hyderabad**                                                                                          *Hyderabad, India*
- Web-System Administrator                                                                            *Fall 2017, Spring 2018*
- TA for Data Warehousing and Data Mining (CSE445), Database Systems (CSE441)                         *Fall 2018, Spring 2019*

# Publications

**SVP-CF: Selection via Proxy for Collaborative Filtering Data**                                   *[LINK]*

NOVEEN SACHDEVA, CAROLE-JEAN WU, JULIAN MCAULEY

In the SubSetML Workshop @ 38th International Conference on Machine Learning – ICML '21

**ECLARE: Extreme Classification with Label Graph Correlations**                                   *[LINK]*

ANSHUL MITTAL, NOVEEN SACHDEVA, SHESHANSH AGRAWAL, SUMEET AGARWAL, PURUSHOTTAM KAR, MANIK VARMA

In The Web Conference 2021 – WWW '21

**Off-policy Bandits with Deficient Support**                                   *[LINK]*

NOVEEN SACHDEVA, YI SU, THORSTEN JOACHIMS

In the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining – KDD '20 (Research Track)

**How Useful are Reviews for Recommendation? A Crit. Review & Potential Improvements**                                   *[LINK]*

NOVEEN SACHDEVA, JULIAN MCAULEY

In the 43rd International ACM Conference on Research and Development in Information Retrieval – SIGIR '20

**Sequential Variational Autoencoders for Collaborative Filtering**                                   *[LINK]*

NOVEEN SACHDEVA, GIUSEPPE MANCO, ETTORE RITACCO, VIKRAM PUDI

In the 12th ACM International Conference on Web Search & Data Mining – WSDM '19

**Attentive Neural Architecture Incorporating Song Features For Music Recommendation**                                   *[LINK]*

NOVEEN SACHDEVA, KARTIK GUPTA, VIKRAM PUDI

In the 12th ACM International Conference on Recommender Systems – RecSys '18

**Explicit Modelling of the Implicit Short Term User Preferences for Music Recommendation**                                   *[LINK]*

KARTIK GUPTA, NOVEEN SACHDEVA, VIKRAM PUDI

In the 40th European Conference on Information Retrieval – ECIR '18

# Achievements

- Reviewer for ICDM' 18, AAAI' 19, KAIS '19,20, ICWSM '21,22
- Received generous travel grant from ACM to present paper at The Web Conference                                   (2021)
- Mentoring a group of three brilliant undergraduates at UC San Diego                                   (2020)
- Awarded the Jacobs School of Engineering Fellowship at UC San Diego                                   (2020)
- Received generous travel grant from SIGIR to present paper at SIGIR                                   (2020)
- Received generous travel grant from SIGIR & Flipkart to present paper at WSDM                                   (2019)
- Qualified for ACM-ICPC Asia Onsite Regionals. *Online: $95^{th}$, Onsite: $120^{th}$*                                   (2018)
- Dean's Research Award for exceptional undergraduate research work at IIIT Hyderabad                                   (2018)
- Dean's Award for Academic Excellence: Top 10% of batch                                   (2018)
- Mentor for Google Code-In at JBoss, RedHat                                   (2017)

# Projects

**LED Display Construction & Home Automation**

INDEPENDENT COVID LOCKDOWN PROJECT                                   *Apr. 2020 - Jul. 2020*

- Built a $42 \times 24$ Wi-Fi controlled LED display using commonly available LED strips. Core features include in-room music visualization using FFT, image cast, animations, etc. Final size equivalent to a $32''$ TV, and able to achieve $60$ FPS.
- Made my own smart switches using electronic relays and micro-controllers. Core features include a small size to fit inside a common switch-board and a significantly lower price than ($10$x) commercial solutions. Bundled everything together with an in-house android application.

**Generative models for Stochastic Point Processes**

INDEPENDENT REMOTE COLLABORATION WITH *Prof. Giuseppe Manco*                                   *Jan. 2019 - Dec. 2019*

- Leveraging the modelling power of weibull distributions to predict user return-time to media-streaming-services like YouTube, Netflix etc.
- Formulating and experimenting with generative models like GANs & VAEs – maximizing the likelihood of the actual return-time.

**PEGASOS: Gradient based solver for SVM**

OPTIMIZATION METHODS COURSE PROJECT UNDER *Prof. C V Jawahar*                                   *Mar. 2019 - May. 2019*

- Implemented both linear and kernelized PEGASOS, results and run-time matching with popular libraries like sklearn.
- In addition to given problem statement, exploited the concept of Gramian Matrices to speed up the learning process for kernelized PEGASOS.

**Compiler for C-like language, Decaf**

*Oct. 2018 - Dec. 2018*

- Created a flex tokenizer and bison parser from a Context Free Grammar, using visitor software design pattern for modular compiler design.
- Parsed source programs into an Abstract Syntax Tree and then into LLVM intermediate representation (LLVM IR).

# Skills

|  |  |
|---:|:---|
| **Languages** | Python, C++, C, Bash, MATLAB |
| **Machine Learning** | PyTorch, Tensorflow, Keras, scikit-learn |
| **Miscellaneous** | LaTeX, Git, SQL, Neo4J, Flask, Node.JS, PHP |