Vishal Keswani

Senior undergraduate, IIT Kanpur

vkeswani@iitk.ac.in | +91 8741 867 137 | Gmail | LinkedIn

EDUCATION

Indian Institute of Technology, Kanpur (May 2020)

- Bachelor of Science, Economics C.P.I. : 8.6/10
- Minors: Machine Learning, Computer Systems, Cognitive Science

Kendriya Vidyalaya 1, Ajmer (May 2015)

- Central Board of Secondary Education Grade: 96.4/100

East Point Senior Secondary School, Ajmer (May 2013)

- Rajasthan Board of Secondary Education Grade: 94.2/100

SCHOLASTIC ACHIEVEMENTS

Scholarships and Grants

- One of 2 Economics students in India to receive prestigious research grant of SURGE, IITK (2018)
- Selected for MCM scholarship by SSPC, IIT Kanpur for one academic year (2016-17)
- Received DONOR scholarship by SSPC, IIT Kanpur for the spring semester (2016-17)

Awards

- Stood 1^{st} in Task A of Memotion Analysis under SemEval-2020 International Workshop (2020)
- Received Commendation Certificate for excellent performance under SURGE'18 program (2018)
- Bagged the Best Staff's Pick Project Award in the course TA201A in a batch of 380 (2018)
- Ranked among top 1.1% out of 0.2 million students in JEE Advanced 2016 with City Rank 2 (2016)
- Secured AIR 538 in JEE Mains 2016 out of 1.5 million candidates with City Rank 1 (2016)
- Awarded Certificate of Appreciation from HRD minister for exceptional performance in C.B.S.E. (2015)
- Received Certificate of Merit for position in the top 1.5% of K.V.S. students in C.B.S.E. examination (2015)

RELEVANT COURSES

- **Computer Science & Electrical Engineering**
  Machine Learning for Signal Processing, Introduction to Machine Learning, Natural Language Processing*, Data Structures and Algorithms, Fundamentals of Computing, Principles of Database Systems*
- **Mathematical Economics**
  Econometrics I (Inference), Econometrics II (Time series, Panel data, Discrete choice), Bayesian Econometrics, Financial Econometrics, Mathematical Economics (Optimization theory)
- **Mathematics & Statistics**
  Probability and Statistics, Stochastic Processes, PDE, Real Analysis & Calculus, Vector Algebra & ODE
- **Cognitive Science**
  Intro. to Psychology & applications, Psychology of language, Philosophy of Cognitive Science*, Social Cognition*
- **Coursera**
  Process Mining (TU/e), Machine Learning (Stanford), Python for everybody (UMich) (*ongoing)

RELEVANT SKILLS

- **Programming Languages** Python, MATLAB, R, C, C++, HTML, CSS, JavaScript, SQL
- **Package Familiarity** Pytorch, Pm4Py, NumPy, Pandas, Nltk, Keras, TextBlob, ScikitLearn
- **Utilities and OS** MS Office, Latex, Bash scripting, Linux, Windows

RESEARCH PROJECTS

Scalable sampling using Variational Autoencoders & Normalizing Flows (Jan'20-Present)

Dr Vipul Arora, Undergraduate Project, Department of Electrical Engineering, IIT Kanpur

- Reviwed Rezende et al. (2014, 2015), Kobyzev ey al. (2019), Kingma et al. (2015, 2018), Berg et al. (2019)
- Implemented VAE and planar flows for approximating complex known densities and MNIST digits
- Currently working on flow-based models (GLOW, Sylvester) for Boltzman distribution sampling

Multimodal Emotion Analysis of memes (Text+Image), SemEval'2020 (Jan'20-Apr'20)

Dr Ashutosh Modi, Department of Computer Science and Engineering, IIT Kanpur

- Reviewed Vaswani et al. (2017), Devlin et al. (2019), Oquab (2014), Keila et al. (2019), Qian et al. (2016) etc
- Implemeted CNN (image-only) and ANN (text-only), combined using SVM, RF and Neural Network
- Fine-tuned BERT (text-only) and Multimodal Bitransformer (text+image), also used Naive Bayes (text-only)
- ANN+Word2vec & Naive bayes perfomred best, beat baseline Macro-F1 by 63%, 39% & 27% for the 3 subtasks

Product Recommendation System (Aug'19-Nov'19)

Dr Purushottam Kar, Department of Computer Science and Engineering, IIT Kanpur

- Explored different algorithms from the extreme multi-label classification repository (manikvarma.org)
- Parabel, PfastreXML and FastXML were the top picks based on performance on most evaluation metrics
- Reviewed Prabhu et al., 2018 (Parabel), Jain et al., 2016 (PfastreXML), Prabhu & Varma, 2014 (FastXML)
- Trained the 3 on a dataset of 10k users, represented by ~16k features, with ~3k labels of interest
- FastXML performed overall best with model size 1.42mb, prec@1 80.4% & prec@3 49.7% on test dataset

Public opinion on US-China trade war through twitter sentiment analysis (Aug'19-Nov'19)

International Economics and Finance, Department of Economic Sciences, IIT Kanpur

- Reviewed Medhat, Hassan, & Korashy (2014), O'Connor et al. (2010), Bermingham & Smeaton (2011) etc
- Extracted twitter data (~10m tweets) filtered by keyword 'tradewar' from Jan 2018 (beginning of trade war)
- Performed data preprocessing - lowercasing, removing stop words etc, & replacing emoticons with words
- Trained Naive Bayes, Maximum Entropy & Decision Tree; NB gave the best accuracy (~74%) & training time
- Classified tweets as positive/negative, classification probabilities gave a polarity score for each tweet
- Aggregated polarity scores to get daily trend, hence, constructed a subjective proxy for policy decisions

ARIMA vs LSTM: Modeling and Forecasting of US real GDP (Jan'19-Apr'19)

Dr Deep Mukherji, Department of Economic Sciences, IIT Kanpur

- Reviewed research in econometric vs deep learning techniques for time series forcasting including
- Ahmed, Atiya, Gayar, El-Shishiny (2010), Siami-Namini, Namin (2018), Brownlee (2016-17)
- Obtained seasonally adjusted real US GDP (quarterly) from the U.S. Bureau of Economic Analysis
- Autocorrelation function, Dickey-Fuller & Phillips-Perron tests supported first differencing for stationariy
- Used 9:1 train-test split due to high prediction time and walk-forward validation (rolling forecasts)
- Used MSE for tuning and model evaluation, ARIMA(1,1,1) performed best among econometric models
- LSTM showed ~23% reduction in test MSE with 2 neurons & 18% with just 1 (batch-size = 1, epochs = 10)

INTERNSHIP PROJECTS

Anomaly detection in telecommunication event logs (May'19-July'19)

Data Science Research Intern, Nokia Solutions and Networks, Gurugram, India

- Explored the research in unstructured processes and unlabeled event logs ranging over below works
- Will van der Aalst (2007, 2008, 2016-book), Ferreira & Gillblad (2009), Vaarandi (2003) and Li (2015)
- Assigned case ids to log data (stream of unlabelled events) using an iterative EM procedure (Ferreira)
- Discovered process using alpha & inductive miner, the resulting petri net revealed a Spaghetti process

- Performed conformance checking and classified event types as anomalous based on aggregate trace fitness score for different thresholds and submitted results for approval by domain expert

Auto-Suggesting inquiry questions based on client case (May'19-July'19)

Data Science Research Intern, Nokia Solutions and Networks, Gurugram, India

- Obtained word documents containing the chat threads between the clients and the care engineers (CE)
- Extracted e-mail bodies to csv, followed by tokenization, removal of stop words, stemming, lower casing
- Used keyword search to create 'queries and CE questions repository', mapped the two with each other
- Vectorized each row using tf-idf scores, utilized for calculation of cosine-similarity measure
- Performed k-means & k-medoids on client queries; for a new query, found the order of similar centroids
- Reported inquiry questions for first 3 clusters (based on correspondence) in decreasing order of similarity

Effect of probabilistic sample size on preference behavior (May'18-Dec'18)

Research Intern, SURGE, Department of Computer Science and Engineering, IIT Kanpur

- Designed a web-based betting game-cum-experiment using HTML, CSS, JavaScript consisting of 4 types of trials corresponding to low and high cognitive sample size in the two settings (hedonic and utilitarian)
- Used the Logistic function to simulate outcomes of bets, applied Box-Muller method to generate a Random Gaussian sample used to calculate percentile of players based on total amount after every bet
- Computed mean reaction times and performed outlier detection using IQR method in MATLAB, utilitarian choices seem to converge for both sample sizes while hedonic ones diverge from each other

OTHER PROJECTS

Non-negative Matrix factorization for Image Classification (Aug'19-Nov'19)

Dr Vipul Arora, Department of Electrical Engineering, IIT Kanpur

- Implemented PCA & NMF from scratch, performed dimensionality reduction on MNIST digits for K=2,3
- Plotted transformed data in 2D and 3D, basis vectors and reconstructed samples for the two cases
- Used NMF for image classification and obtained an accuracy of around ~80% on test data for K=3
- Also implemented Neural Network, GMM and Regression variants in the course from scratch

Corruption and the Indian Economy: Story of the States (Jan'19-Apr'19)

Dr Debayan Pakrashi, Department of Economic Sciences, IIT Kanpur

- Obtained statewise data for the corruption index (% of direct incidents of bribery) from CMS-ICS reports
- The F-test and LM test favoured use of pooled regression over Fixed and Random effect models resp.
- Corruption was found to suppresses economic growth and foster capital expenditure and FDI

Incorporating unproductive activities in the empirics of Economic Growth (Aug'18-Nov'18)

Undergraduate Project, Department of Economic Sciences, IIT Kanpur

- Formed the benchmark regression model (Mankiw, Romer, Weil) with adjusted R-squared of 0.66
- Obtained indices for corruption, terrorism, & crime from Weil lab data (2012), augmented one at a time
- The adjusted R-squared improved to 0.81, 0.74, and 0.79 respectively for the abone three variables

POSITIONS OF RESPONSIBILITY

Company Coordinator, Student Placement Office, IIT Kanpur (Aug'17 - Apr'18)

- Ensured smooth conduction of internship and placement processes throughout the season
- Organized PPTs, GDs, invigilated during tests, interacted with company officials & managed interviews

Senior Executive Media & Publicity, E-cell, IIT Kanpur (Apr'17 - Apr'18)

- Created articles and posts for online media for promotion of various events & the E-cell of IIT Kanpur
- Aided in organization of eSummit and TEDx, interviewed Mr. Sandeep Aggarwal (founder of ShopClues)

TEAM PROJECTS

Effect of Mood congruency during encoding/retrieval on recall PSY152 (May'17 to Jul'17)

- A group of 6 designed & conducted an experiment relating memory & mood during encoding & retrieval
- Established a link between congruent mood states & ability to recall neutral words, got **A*** in the course

Impact of emotion inducing stimuli on selective attention PSY151 (Jul'16 to Nov'16)

- Coordinated a group of 6, developed an experiment using emotion inducing stimuli (pictures, videos)
- Stimuli triggered problem-solving ability & attention in a specific task, **1st** position in a class of 150

Rubber Band Machine Gun TA201 (Jan'18 to Apr'18)

- Spearheaded a group of 5 batchmates to design a metal gun that fired 120 rubber-bands in a few sec.
- Used brazing, sheet-metal forming, casting & soldering; built a visually appealing hand-driven structure

Mechanical Tower Crane TA202 (July'17 to Nov'17)

- Worked in a group of 6, fabricated a crane using turning, milling, drilling, ideated without Google help
- Could lift as well as transfer loads in a cylindrical span determined by its adjustable height and radius

EXTRA-CURRICULAR ACTIVITIES

- Secured 1st position in British Parliamentary Debate in Galaxy (Inter-pool cultural fest of IITK) (2017)
- Worked as Secretary in Fine Arts Club and in Ritambhara, the agship fashion event of Antaragni (2017)
- Anchored in the department Freshers' Night, also performed speed art in batch Freshers' night (2016)
- Became School Captain of Kendriya Vidyalaya 1, Ajmer, also won Best Student award (2015)
- Best Speaker from region in Youth Parliament Competition, Zonal level (5 zones in India) (2013-14)
- Captained School Cricket Team and won Fair-Play award in an open cricket tournament (2013)
- Stood in top 150 in Tata Building India Essay Competition among 2m participants twice (2010-11 & 11-12)