Deepali Jain

Big-Data Experience Lab
Adobe Systems
Bangalore, India - 560087

Email : jaindeepali811@gmail.com
Website : jaindeepali.github.io
Phone : +91 826-680-1703

RESEARCH INTERESTS

Reinforcement Learning, Sequential Decision Making, Knowledge Representation and Reasoning

EDUCATION

Indian Institute of Technology Roorkee *Roorkee, India*
Bachelor of Technology in Electrical Engineering *2012 – 2016*

RESEARCH/INDUSTRY EXPERIENCE

Adobe Research, Big-Data Experience Lab, Research Fellow *June 2016 - Present*
Adobe Research, Big-Data Experience Lab, Research Intern *May 2015 - July 2015*
Google Summer of Code, Wikimedia, Open-Source Developer *May 2014 - August 2014*

PUBLICATIONS

Measurement of Users' Experience on Online Platforms from their Behavior Logs
D. Jain, A. Sinha, N. Sheoran, D. Gupta, S. Khosla *(Submitted to PAKDD 2018)*

An LSTM Based System for Prediction of Human Activities with Durations
K. Krishna*, **D. Jain***, S. Mehta, S. Chaudhary *(Accepted PACM-IMWUT, UbiComp 2018)* (* Equal contribution)

Fairness Aware Recommendations on Behance
N. Modani, **D. Jain**, U. Soni, G. K. Gupta, P. Agarwal *(PAKDD 2017)*

PATENTS

Predictive Analysis of User Behavior utilizing RNN-based User Embedding
S. Kim, C. Chen, H. Bui, E. Koh, B. Kveton, **D. Jain**, A. Sinha, D. Gupta, N. Sheoran *(US Patent App. 15/814,979. Filed 17 November 2017)*

Creator Aware and Diverse Recommendations of Digital Content
N. Modani, **D. Jain**, U. Soni, G. K. Gupta, P. Agarwal *(US Patent App. 15/598,193. Filed 16 June 2017)*

Personalized Creator Recommendations
N. Modani, **D. Jain**, U. Soni, G. K. Gupta, P. Agarwal *(US Patent App. 15/625,237. Filed 17 May 2017)*

RESEARCH PROJECTS
SEQUENTIAL DECISION MAKING

Measurement of Users' Experience on Online Platforms from their Behavior Logs *Under review*
Users' behavior on an online platform is modeled as a partially observed Markov decision process (POMDP). User experience (UX) is measured as a latent variable represented by the value function in the POMDP.

- Implemented environment simulation using next action prediction model. Average F1 score: 47%.
- A function approximation based fixed point iteration method is implemented to compute UX.
- Formalized another rule-based approach informed by consumer psychology as a benchmark.
- Evaluation of the method provides evidence that estimated values are relevant metrics of UX (69% F1 score on conversion prediction). The proposed approach can overcome deficiencies of traditional survey methods for experience measurement.

An LSTM Based System for Prediction of Human Activities with Duration *Published*
Two RNN network architectures for joint prediction of human activities and their duration are proposed.

- Designed a activity data collection experiment and created an android app for the same.
- Designed a Recurrent Neural Network (RNN) architecture that shows 20% higher accuracy (47.6% on 21 class classification) than the best baseline.
- Developed a probabilistic sequence matching algorithm and implemented a string alignment algorithm inspired from gene and protein sequence mining research as comparison baselines.
- Implemented and compared methods of sequence generation from RNN for modeling complete days.

- Designing a multi-goal RL agent to suggest changes to the daily routine of a subject. Objective is to achieve certain productivity goals with minimum disruption in regular activities.

An Auction Based Mechanism for Recommending Items in a Two-Sided Platform *Working paper*

An auction based mechanism is proposed for recommendation in multi-stakeholder setting. The objective is to ensure relevance of content to viewers and to increase satisfaction of content owners by allowing them to target desired audience.

- Supervised a team of interns for implementing a simulation of an online auction system for recommendation of content from 650 creators to 1500 viewers.
- Designed and carried out experiments to show that while, as one can expect, the relevance of the recommendations to the viewers reduces slightly, the proposed system is able to achieve the objectives of the content-owners significantly (170%) better.
- Designing an online learning algorithm to adapt bidding strategy in response to changing market value by estimating true value of other bidders.

TEMPORAL MODELING/ SIGNAL PROCESSING

Emotion Detection from Sensor Data of Wearable Devices *Working paper*

- Conducted a field experiment for inducing emotions with varying valence and arousal levels in subjects while monitoring their heart-rate.
- Tested the hypothesis that human emotions can be detected from heart-rate signal from a wearable fitness device.
- Designed frequency domain features through Short Term Fourier Transform and Discrete Wavelet Transform to capture heart-rate variability.
- Implemented an RNN sequence classification model and a deep Convolutional Neural Network (CNN) for recognition of emotion from heart-rate signal (Average F1 score: 51% on 6 class classification).

Blind Source Separation of Audio Signals *Tech Report, May 2016*

BSS is the problem of separating mixed signals from unknown sources.

- Conducted a comparative study of the existing higher-order statistical signal processing solutions to to the Blind Source Separation (Cocktail Party) problem such as Independent Component Analysis, Singular Value Decomposition and Multiple De-correlation.
- The accuracy as well as efficiency results obtained from same experiments on different methods were documented, showing limitation of existing methods in separation of real-world audio signals due to convolutive effect of sound reverberation.

OTHER RESEARCH

Representation of Marketing Domain Knowledge *Ongoing*

- Created a prototype to represent textual knowledge from Marketing literature in the form of a knowledge graph and draw first-order inference on it in Prolog for a question-answering application.
- Compared existing memory-augmented and attentive neural network architectures for the task of question answering from text.

Fairness Aware Recommendations on Behance *Published*

- A re-ranking strategy was proposed that can be applied to the scored recommendation lists to improve exposure distribution across the content creators on two-sided platforms (thereby improving the fairness), without unduly affecting the relevance of recommendations provided to the consumers.
- Implemented a scalable spark app for item similarity based collaborative filtering algorithm. Also implemented a greedy optimization based reranking algorithm for making the collaborative filtering based recommendations fair and diverse.
- Carried out experimentation to show that our method resulted in recommendations with much higher level of fairness (70%) and representative diversity (130%) compared to the state-of-art recommendation strategies, without compromising the relevance score too much.

SELECTED PROJECTS

Human-Computer Interface using Electroencephalography Signal Classification *Aug - Nov 2015*

Recorded EEG signals were pre-processed through Common Spatial Pattern filter and a binary SVM classification model was trained on the samples to predict motor imagery classes (left/right hand movement).

Context-Aware Media Content Analysis *May - Jul 2015*
Built an end-to-end physical activity recognition system based on accelerometer data from mobile sensors. Tested the hypothesis that different context of readers leads to different choice of media articles, using click-stream data analysis.

Adler: Text Classification API based on TechTC-300 Test Collection *Feb - Apr 2016*
Developed a ready-to-use text corpus generation engine as an open source python package. The final dataset used chi-squared feature selection and TF-IDF feature weighting. Classification of text into various topics based on extracted features was performed using a Decision-Jungle classifier in AzureML Studio.

Raphael: Classification of Paintings based on Style and Era *Apr - May 2015*
Artificial Neural Networks Course Project under Prof. G.N. Pillai. Built a system to classify paintings taken from wikiart.com into the era and style they belong to, using low level (color histograms, texture, color saturation) and high level (Scale-Invariant Feature Transform (SIFT) descriptors, Speeded-Up Robust Features (SURF)) image features. The extracted features were quantized using K-means clustering and bag-of-visual-words method. An ensemble of SVM, Feed-forward Neural Networks and Random Forests were used for prediction of painting style and era.

Forsit: Hybrid Recommendation System for Mathematical Problem-Solving Website *Aug 2015*
Winning entry for Ideaz Paper Presentation contest. Built a recommendation engine for a mathematical problem-solving platform developed by SDSLabs using a hybrid of content based and collaborative filtering methods.

Xavier: Knowledge Representation of PDF Document Content *Sep 2015*
Built a python module for extracting structured information from a set of PDF documents. The documents were scraped and their content was parsed to extract the document structure as the first step. Then, important entities and relations between them were identified with the help of structure information and Entity Recognition.

Book Management Software for Wikibooks *May - Aug 2014*
Google Summer of Code project. Worked on a PHP and Javascript based extension, BookManager, for Mediawiki software under Wikimedia organization. Created a robust and user-friendly interface for editing, reading, navigation and migration of large Wikibooks ($\sim 10,000$ sections).

Awards and Achievements

Winner, Microsoft Code.Fun.Do, IIT Roorkee *2015*
National Finalist, GSQuantify Data Science Challenge, Goldman Sachs *2015*
Winner, Ideaz Paper Presentation Contest, IIT Roorkee *2015*
National Winner, IEEE Programming League *2015*
Air Cmde S.C. Mehra Scholar, IIT Roorkee *2012*
Best academic performance by woman student in Electrical Engineering
Kishore Vaigyanik Protsahan Yojana Fellow *2012*
High school student with talent and research aptitude for higher studies
National Talent Search Scholar *2008*

Relevant Coursework

Undergraduate: Linear Algebra, Multivariate Calculus, Partial Differential Equations, Transform Theory, Numerical Methods, Artificial Neural Networks, Discrete Mathematics, Mathematical Modeling, Digital Signal Processing
Independent: Reinforcement Learning (Udacity), Deep Reinforcement Learning (UCB), Probabilistic Graphical Models (Coursera), Game Theory (Coursera), Introduction to Machine Learning (Coursera), Deep Learning (Udacity), Knowledge-Based AI: Cognitive Systems (Udacity), Natural Language Processing with Deep Learning (Stanford)

Technical Skills

Machine Learning: TensorFlow, Keras, Spark, Python Scipy stack, R, Theano, Torch, Caffe, GNU Octave, MATLAB
Programming Languages: Python, C++, Java, PHP, JavaScript
Misc.: Bash, Git, Android Development, ReactJS, NodeJS, Django, SQL, LaTeX

Positions of Responsibility

Developer, SDSLabs, IIT Roorkee *2013 - 2016*
Coordinator, Programming and Algorithms Group, IIT Roorkee *2014*
Chief Web Coordinator, Watch Out News Agency, IIT Roorkee *2014*

References

Prof. G.N. Pillai *Professor, IIT Roorkee*
Dr. Shiv Saini *Sr. Research Scientist, Adobe Research*
Prof. Atanu R. Sinha *(Emeritus, University of Colorado Boulder)* *Principal Scientist, Adobe Research*