

Amruta Purandare

Address: Electronic City - Phase 1, Bangalore 560100
Tel: +91 733-843-8215 (Mobile) || +91 880-512-0109 (WhatsApp)
Email: amruta@smartv.in



Online Profiles:

- ❖ Personal Website: <http://www.pamruta.com>
- ❖ LinkedIn Profile: <https://www.linkedin.com/in/pamruta>
- ❖ Code Samples on Git Hub: <https://github.com/pamruta>

Areas of Interests:

[Artificial Intelligence \(AI\)](#) | [Machine Learning](#) | [Computer Vision](#) | [Video Analytics](#) | [Image Processing](#) | [Wearable Devices](#) | [Multimedia](#) | [Digital Entertainment](#) | [Robotics](#)

Technical Skills:

[Python](#) | [Jupyter Notebook](#) | [Amazon Web Services \(AWS\)](#) | [Alexa](#) | [Google Cloud](#) | [Tensorflow](#) | [Pytorch](#) | [Keras](#) | [Kibana](#) | [Elastic Cloud](#) | [Solr](#) | [Raspberry Pi](#)

Education:

Bachelors in Computer Engineering (1998-2002)
Pune University, India
First Class with Distinction - Best Outgoing Student

Masters in Computer Science (2002-2004)
University of Minnesota, USA
Research fully sponsored by National Science Foundation (NSF)

Internships:

SONY Corporation (Tokyo, Japan)
Position: Summer Intern in Intelligent Systems Research Laboratory (ISRL)
Project Areas: Information Extraction, Web Mining, Recommendation Systems
Duration: 3 months (summer 2008)

Amazon.com (Seattle, Washington)
Position: Summer Intern in Software Development and Engineering (SDE)
Project Areas: Data Mining, Machine Learning, e-Commerce, Product Ontology
Duration: 4 months (summer 2005)

Entrepreneurship:

[SMARTV.in](#) (Jan 2010 - Mar 2015) || [CINEMATRIX.in](#) (Apr 2019 - Present)

Domain: **Multimedia & Entertainment Technology Start-Up**
Building the next generation of creative media and digital entertainment technology through AI driven machine intelligence and advanced audio-video & image analytics

Recognized under Start-Up India Scheme & NVIDIA Inception Program

Work Experience:

Associate Director at Happiest Minds Technologies (Apr 2017 - Oct 2018)
Division: Analytics Center of Excellence (CoE)
Project Areas: Computer Vision, Image Processing, Deep Learning, Video Analytics

Principal Data Scientist at MakeMyTrip (Aug 2016 - Mar 2017)
Division: Technology Development (Bangalore, India)
Project Areas: Text Mining from Hotel Reviews and Wiki-Travel Pages

Project Lead at Persistent Systems Ltd (Apr 2015 - May 2016)
Division: Business Intelligence & Data Analytics (Pune, India)
Project Areas: Data Mining, Machine Learning, Predictive Analytics

Academic Research Positions:

Research Staff - Singapore Management University (Singapore)
Division: Information Systems
Project Areas: Social Network Analysis, Data Mining, User Behavior Modeling
Research sponsored by Defense Science Organization (DSO)
Duration: 1 yr (2009)

Research Assistant - University of Pittsburgh
Division: Intelligent Systems Program
Areas: Dialog Understanding, Conversational Interfaces, Spoken Language Technology
Research sponsored by Office of Naval Research (ONR)
Duration: 4 yrs (2004-2008)

Visiting Research Student - University of Southern California (USC)
Division: Information Sciences Institute (ISI)
Project Areas: Information Extraction, Lexical Semantics, Knowledge Discovery
Research sponsored by Department of Homeland Security (DHS)
Duration: 3 months (summer 2007)

Research Assistant - University of Minnesota, Duluth
Department: Computer Science
Project Areas: Natural Language Understanding, Text Clustering, Word Semantics
Research sponsored by National Science Foundation (NSF)
Duration: 2 yrs (2002-2004)

Awards and Achievements:

Invited speaker at Indian Technology Congress, 2019

Interview with [Analytics India Magazine](#) on the occasion of Women's Day 2019

“Technology Innovation Award” - by Happiest Minds Technologies (Dec 2017)

“You Made a Difference Award” - by Persistent Systems Ltd (Nov 2015, Mar 2016)

“Best Outgoing Student” - by Cummins College of Engineering, Pune (2002)

“India Foundation Scholarship” - by Sakal Media Group (2002)

“Best Paper in Artificial Intelligence & Fuzzy Logic” - by Pune Institute of Computer Technology (PICT), in association with IEEE (2001)

“People's Choice Award” - by University of Pittsburgh (2006)

Project Summary:

1. Demo Prototypes built under the start-up initiative **CineMatrix**:
 - a. Using **OpenPose** library to capture body movements and gestures in Dance & Figure Skating Videos
 - b. **Colorization & Digital Restoration** of Old Black & White Films using **DeOldify**
 - c. **Automatic Scene Generation** using DC-GANs in PyTorch. Model is trained on Intel Image Classification Dataset that contains images of scenes from glaciers, mountains, forests, streets, buildings etc.
 - d. **Neural Style Transfer** to create artistic photo filters using Keras and PyTorch.
 - e. Identifying animal species in wild-life films using **YOLO V3 Object Detection**
2. Presented AI / ML demos on Audio, Video and Image Analytics to senior executives at ESPN on the following topics:
 - a. **Sports Activity Recognition**: Classifies images by recognizing sports activities (e.g. swimming, cycling, skating, boxing, surfing etc) using **Transfer Learning**.
 - b. **Video Tagging & Meta-Data Extraction**: Extracts meta-data from sports videos using techniques like Optical Character Recognition, Speech-to-Text and Face Recognition, to identify players, teams, sports, tournament and venue details.
 - c. **Contextual Advertising**: Displays product ads for sports merchandise (e.g. boxing gloves, wetsuits, tennis racquet, roller-skates) based on the activities recognized in sports videos.

3. Mentored a team of 7-8 interns from **BITS Pilani** on the following AI / ML Projects:

- a. **Product Image Search & Visual Similarity:** Given an input image, searches the product catalog for visually similar items using Deep Learning library, **PyTorch**. The algorithm is applied to various product categories: clothing (shirts, skirts, jackets, dresses etc), home decor (curtains, bedsheets, lamps, carpets) and fashion accessories (shoes, watches, jewelry, handbags etc).
 - b. **Contextual Advertising:** Recognizes objects in the given video stream, and searches the product catalog using the above algorithm to find similar items. Product ads for similar items are embedded in the video using overlay. This concept is applied to Fashion Show videos, to find dresses and clothes similar to those worn by fashion models on the screen.
 - c. **Sports Video Analytics:** Automatically generates highlights for sports videos using audio-visual features that capture the excitement & action in the game. Also, face recognition is used on close-up shots to automatically display career highlights or brief bio of players on the screen.
 - d. **Entity-Relation Graph:** Displays relations between sports players and the teams they play for, as well as, between the teams and the leagues they are part of, using the text extracted from Wikipedia. Stanford's **Core-NLP** library is used for parsing text to find entity-relations, and **D3.JS** for creating visualizations.
 - e. **Smart Cities:** Analyzes video streams from traffic cameras to count the vehicle flow (number of cars, trucks, bikes and buses passing) on the street. The same algorithm is also extended to create a **Smart Parking** app, which captures an image of a parking lot, and creates a map of empty and full parking lots, to guide users to the nearest empty parking lot.
4. Guided a team of 6-8 data scientists at **Happiest Minds Technologies** on the following AI / Machine Learning Projects:
- a. **Smart Kiosk** - solution developed for **retail** outlets. Customers scan products at Smart Kiosk to see product information, like its features and specifications, nutritional facts and recipes for food items, and related products for cross-selling. Smart Kiosk uses **Computer Vision** to recognize fresh produce like fruits and vegetables, and **Optical Character Recognition (OCR)** to read product labels, brand names and logos.
 - b. **Video Highlights Generation** - The algorithm uses **audio signal processing** to detect loud cheering, clapping and applause for automatically generating highlights for **sports videos**.
 - c. **Damage Detection** - This **deep-learning** based **image classifier** uses TensorFlow model to identify signs of **visible damage like scratches and cracks** on mobile phone screens, car windshields and glass windows to speed up the process of **insurance claims processing**.

- d. **Audio Classification** - Classifies **ambiance sounds** by detecting car honking, dog barking, clapping, screaming, door knock etc using **PyAudio** library.
 - e. **Health Bot** - A conversational chat-bot to search nearby doctors and hospitals by specialty areas like cardiologists, dentists, pediatricians etc.
 - f. **Video Analytics** - Face Recognition, Person Tracking, Vehicle Number Plate Recognition for CCTV surveillance cameras.
 - g. **Video Tagger** - Generates audio transcripts by running **Speech Recognition** on TEDx talks, online courses and lecture videos to identify important topics.
5. Prepared a complete roadmap for AI & Machine Learning group at Happiest Minds presenting solutions and offerings around **Audio, Video, Image and Edge Analytics**.
6. Projects on **Raspberry Pi, Arduino, Alexa & Wearable Devices** -
- a. **Media player and Car Infotainment System** built on Raspberry Pi.
 - b. **SPiCam**: Video Surveillance camera for Smart Homes.
 - c. **Smart Fridge**: Recognizes items in the fridge using the AWS Rekognition library.
 - d. **Smart Jewelry**: Using Near Field Communication (NFC) Technology.
 - e. **FairyTell Bot**: Alexa skill to read bedtime stories from children's classics.

These projects can be found on **Hackster** at: <https://www.hackster.io/pamruta>

7. Created a **Kibana dashboard** to visualize **crowd-funded projects on Kick Starter**. Projects can be visualized by themes, categories, funding received and launch date.
8. Built a **culinary search-engine** to search thousands of **online recipes** based on the given **ingredients** (e.g. *rice, lentils, spinach, corn, mushroom*) or **category** (e.g. *sandwich, soup, cake, breakfast, brunch, vegetarian*). Sample search-queries include: *"apple pie", "carrot cake", "mushroom soup", "potato salad"* etc.
9. Projects done in **MakeMyTrip** -
- a. Built a **destination-search engine** by mining articles on **Wiki-Travel**, to suggest top domestic and international cities for a selected **activity** (e.g. scuba diving, ice skating, cross-country skiing, trekking, horse riding etc) or **theme** (e.g. wild life safari, hill station, amusement parks, world heritage sites etc).
 - b. Built a **hotel-search engine** by mining reviews from **Trip Advisor**. Hotels can be searched near specific **points-of-interests** (POIs) like *metro station, airport, popular landmarks, local attractions* and *neighborhoods*, or by specialty services like *Italian Dining, Infinity Pool, Private Beach* etc.

- c. Performed **sentiment analysis on text-snippets (phrases)** extracted from online hotel reviews to **score and rank hotels** on various dimensions like location, quality of food, service, cleanliness and amenities.
- d. Trained a **word2vec model** on Trip Advisor reviews to automatically **discover concepts** related to travel domain from text. Also, created **visualizations** by projecting word vectors in 3D space using the **Embedding Projector** visualization tool from Google.

10. Projects done in Persistent Systems Ltd -

- a. **Automatically generated skill-profiles** for technical support engineers by mining unstructured text-content from email communications. The algorithm retrieves a **ranked list of experts in the given skill-areas**.
- b. Built a demo prototype for a client that works in the medical and health-care domain by **analyzing the content posted on social media**. The project identifies **trending topics, user activities**, as well as, the **sentiments** in user posts, to study the **effects on product sales and revenue**.
- c. Performed **predictive analysis** for a client that offers bike rental service in the Bay Area, by computing **correlations between bikes rented and weather parameters** like temperature, wind speed, humidity etc. on a given day.
- d. Implemented a **brand-clustering algorithm** to identify companies that work in the **same industry sectors or offer similar products and services**, by extracting features from online news text using **word2vec** utility.
- e. Mentored a team of college interns from the College of Engineering, Pune (COEP) on **“Box-Office Predictions”** project, to **estimate the box-office revenue** for movies, based on parameters like reviews & ratings, production budget, studio, cast and genres.

Won **“You Made a Difference”** award for contributions on the above projects.

11. Projects at University of Pittsburgh -

- a. Proposed a novel technique for detecting and analyzing **humor in comedy television show FRIENDS**. This research was published in one of the top international conferences in Natural Language Processing area (**EMNLP**) held in Sydney, Australia in 2006. The paper analyzed **dialog transcripts and audio recordings** in FRIENDS TV-show for **automatic humor detection**.
- b. Proposed a Machine Learning framework for **analyzing coherence in spoken conversations**. The algorithm tries to **distinguish random incoherent conversations from natural coherent dialogs** with over 85% accuracy. This research was published in **FLAIRS 2008 conference** held in Florida.

12. Developed classification algorithms for predicting the email reply order for **Automatic Email Prioritization**. The project analyzed user behavior and inter-personal relationships among users, along with the features extracted from emails to **predict the email-reply order** for prioritization.

13. **Summer Internship Projects -**

- a. Implemented Text Mining and Information Extraction algorithms for **automatically building a large-scale relational database** for movie actors and pop-singers by **mining web pages and biographies on Wikipedia**. This project was carried out at **SONY Corporation, Japan** during summer 2008 internship.
- b. Explored data mining techniques to **automatically categorize product items** based on the similarity of their features and product descriptions. This project was done at **Amazon.com, Seattle** during summer 2005 internship program. The project utilized **unsupervised clustering** algorithms to automatically build a **product taxonomy by grouping similar products**.
- c. Implemented language semantics demo for analyzing **similarities and relations between entities from natural language texts**. This project was done at the **Information Sciences Institute (ISI) of University of Southern California (USC)** during summer 2007 internship.

14. **Projects at University of Minnesota -**

- a. Developed an open source software package **SenseClusters** for **Unsupervised Word Sense Discrimination** task. This project was funded by the **National Science Foundation (NSF)** research grant and was completed as part of the **Master's Thesis at University of Minnesota**.
- b. Worked with collaborators at **Mayo Clinic** in Rochester, Minnesota on bioinformatics project to develop **unsupervised learning methods for resolving ambiguities in biomedical texts**.

15. Won 2nd Prize for **Best Paper in Artificial Intelligence & Fuzzy Logic** in technical/research paper presentation competition organized by **Pune Institute of Computer Technology (PICT), India** in association with **IEEE**. The paper presented a prototype model for a **language understanding system using parsing and logical inference**.

16. Worked on a **Contextual Advertising** project for a start-up company, **Social Extract**. The project analyzes **text content on Twitter** to identify twitter users and relevant tweets for **advertising and marketing campaigns**.