|  |  |
| --- | --- |
| Puneetha Pai  **ML Engineer/Data Scientist, BTech (SJCE, VTU), MTech (BITS Pilani)** | (+91) 8277653959  [puneethapai29@gmail.com](mailto:puneethapai29@gmail.com)  [LinkedIn](https://www.linkedin.com/in/puneeth-pai-b3b299a1/)  [github.com](https://github.com/PuneethaPai) |
| SUMMARY Data and ML enthusiast who transitioned into DS/ML engineer role from Application Developer. Keen on practical and hands-on Machine Learning with proper understanding of theoretical concepts. Believe in mentoring and sharing as a way to grow. Contributing to opensource and part of DS community.  Been in industry for 4 years and worked at various levels of data science sophistication with differing responsibilities. Open Source Contributions:  * [Python-igraph](https://igraph.org/python/): Collection of graph and networks analysis tools * [DVC](https://github.com/iterative/dvc): Data Version Control for ML projects * [Pandas](https://github.com/pandas-dev/pandas): Data analysis and manipulation tool * [EmoPy](https://github.com/thoughtworksarts/EmoPy): A deep neural net toolkit for emotion analysis via Facial Expression Recognition (FER)  EXPERIENCE **ThoughtWorks - *AI-Studio*** Dec 2018 - PRESENT  **Skill Ontology:**   * Identify Skill to Skill relation using Market Basket Analysis over resume and CV data * Build a Skill-People graph to identify people with related skills for staffing needs. * Skill and learning recommendations for people.   Technologies Used: Neo4j, Graph Algorithms, Apriori, DVC, Python  **Superlative:**   * Identify indirect mentions of organization for Redaction Use case. * Approach 1: Build a text classifier using Cosine Similarity to identify Superlative Phrases. * Approach 2: Use POS tagger to identify superlative adjectives. Then customize dependency parser to classify business centric superlative phrases.   Technologies Used: spaCy, Scikit-learn, similarity measure, python  **NER:**   * Sanitize engagement document using Named Entity Recognition. * Deploy a labeling solution, Prodigy, with client domain data for creation of labelled data set. * Automate complete workflow of pre-process, train and model validation using DVC pipelines.   Technologies Used: spaCy, Prodigy, DVC, Python  **ThoughtWorks – *University Trainer*** Apr 2018 - Nov 2018   * TW-University is a grad onboarding/training program to teach Software Development and Agile best practices by simulating a client project and sessions. * I lead a team of 15 trainees, closely mentored 4 of them. * Took sessions on practices like Agile Software Delivery, Consulting. * Gave technical session on test strategy, blue green deployment, feature toggles, etc.   Skills: Teaching, Coaching, Mentoring, Counselling  **ThoughtWorks – *Data Practices India*** Dec 2016 – March 2018  **Inventory Management:**   * Train CNN deep learning model for object recognition. * Built a custom video processor to remove noise (e.g.: customers, staff) from inventory/store to focus on shelves, thus making it easier for prediction by the model.   **Chat Bot:**   * Built a chatbot for Trainline for getting schedule and booking tickets. Integrated it with Facebook messenger interface   **Reinforcement Learning:**   * Implemented custom version of genetic algorithm for an optimization problem   Technologies Used: Keras, CNN, Genetic Algorithm, Python, api.ai EDUCATIONBITS Pilani WILP — *MTech* Software Systems with Specialization in DS and ML  CGPA: 9.7 SJCE Mysuru — *BTech* Electronics and Communications  CGPA: 9.37 | TOOLS **Python**, **DVC, Scikit-learn, Neo4j, AWS,** Jupyter Notebooks, SQL, Hive/Hadoop, Git etc. TECHNIQUES Machine Learning, Statistics, Data Visualization, NLP, Bayesian Methods, Neural Networks, Graph Theory, etc. TRAITS Very **quick technology uptake** due to wide exposure. Loves science, mathematics, and art. Communicates **clearly and concisely**.  Reads, listens, and watches for **continuous improvement** in hard skills, soft skills, and work processes. LANGUAGES  * English * Kannada * Hindi * Konkani |