331 Harbor Way
Ann Arbor, MI

SANSITHA NANDAKUMAR

sansitha@umich.edu
734-773-5925

PROFESSIONAL SUMMARY

An applied data scientist passionate about developing state-of-the-art data science and analytics solutions to solve real world problems. Competent developer that can build end-to-end data science pipelines (ETL -> Feature Engineering -> Modeling -> Deployment)

EDUCATION

University of Michigan, Ann Arbor, MI

— Master's in Information
GPA: 3.88/4 Expected Dec 2020

— Master's in Sports Management
GPA: 3.88/4 Expected Dec 2020

Virginia Tech, Blacksburg, VA

— B.A. in Economics
GPA: 3.34/4 Dec 2017

RELEVANT COURSEWORK

Programming I & II (Python), Data Mining & Manipulation, Information Retrieval & Viz, Intro to Statistics, NLP Algorithms & People, Building Interactive Apps, Needs Assessment & Usability, Agile Development

SKILLS

— **ML Algorithms**: Regression (linear, lasso and ridge), Classification (logistic, svm, decision trees), Ensembles (Random forests, Gradient boosting) and Deep Learning (tensorflow, pytorch)

— **Languages**: Python (tensorflow, pytorch, sklearn, pandas, numpy, pyspark), R and SQL

— **Source Control**: Git

— **Front End**: HTML, Django, CSS

— **Visualization**: Tableau, Domo, Google Analytics

ACHIEVEMENTS

— Varsity tennis player at Virginia Tech
— Won multiple national and international tennis tournaments

EXPERIENCE

— **Research Assistant**
School of Information, University of Michigan May 2020 – current

- Developed an image classifier to detect whether a given Kente cloth (produced in Ghana) is authentic or fake. Extracted image features using Resnet50, MobileNetV2 models and trained a logistic regression classifier
- Deployed the model as a flask-based web app where users can upload an image and get the classification result immediately. Model achieved an accuracy of 88% across an even split of real and fake cloths
- Built a pattern recognition algorithm to identify symbols in a given Kente cloth image using local binary patterns with a decision tree classifier

— **Graduate Student Instructor**
School of Information, University of Michigan Aug 2020 – current

- Student instructor for data-oriented programming (Python). Created assignments, led discussion sections, and hosted office hours

— **Data Analytics Intern**
Bright Spot Solutions, TN, India June 2019 – Aug 2019

- Built a data pipeline with SQL, Apache Spark & Python to compute product performance metrics for any given time bucket.
- Metrics included total revenue, customer details, number of sales, discounts offered, etc.
- This helped in identifying target customers and deciding future promotions

PUBLICATIONS

— Robinson K.P., Eglash R., Bennett A., Sansitha N., Lionel R. **Authente-Kente: enabling authentication for artisanal economies with deep learning**. *AI & Soc* (2020). https://doi.org/10.1007/s00146-020-01055-2

PROJECTS

— **MCity Connected Car**: Mined sensor data from human driven cars to extract insights on driving characteristics under conditions such as stop sign, weather, football game, etc. (SQL, Hadoop and Python)

— **NLP Question Answering:** Achieved an F1 score of 0.81 on the Stanford SQUAD dataset (answer questions given a context) by pairing a pre-trained BERT model (for features) with a pytorch neural network

— **Usability study of Google Voice Assistant:** Collaborated with the UX team at Google to conduct surveys, heuristic evaluation, quantitative analyses, and usability tests to understand what young adults love and hate about using Google Assistant app on Android/iOS. Gave recommendations to improve the app in areas such as privacy control, returning concise answers and using shopping lists