Ryan Muraglia

Address 7249 York St, Dexter, MI 48130
Phone (310) 940 8613
Email rmuraglia@gmail.com
LinkedIn /in/ryan-muraglia

SUMMARY

Data scientist with 5 years of experience developing, deploying and maintaining machine learning models with terabyte-scale streams of big data. Proven track record of delivering business value with both classic data science and analytics, as well as with modern deep learning methods. Strong communicator excels in leadership roles, but also flourishes in independent and collaborative technical roles. Passionate about establishing best practices and processes for rigorous and reproducible science.

WORK EXPERIENCE

Verizon Media (formerly Yahoo!, Oath) — Champaign, IL
Data Scientist II, Brand Data team lead — April 2019 - Present

· Oversaw strategy, development and socialization of our internal *Brand Data* platform, leveraged by partner teams to generate **$200M** in quarterly revenue (influenced and indirect)
· Developed a **PySpark** pipeline implementing a state of the art **deep learning** model in **PyTorch** for named entity recognition (NER), processing over **100 million records/day**
· Created an active learning pipeline to continuously obtain new labeled data, and **automatically retrain and deploy** the NER model to combat model staleness and concept drift
· Prototyped applications of *Brand Data* for adoption by partner teams in sales and targeting, including customer segmentation and churn analysis (**clustering**), brand search topic association (**topic modeling**), and polarized content engagement (**sentiment analysis**)
· Demonstrated a leading and innovative spirit, with a **100% success rate** (3/3 years) in accepted submissions to Verizon Media's annual internal technical conference
· Engaged with the broader data science team to shape policies, processes and best practices for developing high quality code and reproducible models, leveraging Git and Jira integrations

Data Scientist I — June 2017 - March 2019
Data Engineer I — January 2017 - May 2017

· Developed a method for canonicalization of item names extracted from purchase receipts, resulting in a 30% reduction in cardinality, facilitating downstream analysis
· Analyzed the adverse effect of excessive ad load on user engagement and lifetime value
· Contributed to improvements and modernization of a demographic prediction model

EDUCATION

Duke University, Durham, NC — August 2012 - August 2016
M.S., Computational Biology & Bioinformatics (CBB)

University of Michigan, Ann Arbor, MI — September 2007 - April 2011
B.S., Microbiology; Academic Minor, Physics

SKILLS AND PROFICIENCIES

Primary Languages	Python, R, SQL (Hive, Presto), Shell (Bash), Spark (PySpark)
Common Packages	Pandas, Scikit, Tensorflow 2, PyTorch, Tidyverse, Hugging Face
Classic Data Science	Regression and Classification (Linear models, Tree models, Ensembles) Clustering (e.g. k-means), Anomaly Detection (esp. time series)
Deep Learning	Feedforward networks, Convolutional networks (CNN), Recurrent networks (LSTM), Attention networks (Transformers)
Dashboarding	Shiny, Flask, Looker
Additional Tools	Git, Hadoop, Make