

Josue Caraballo

josue.caraballo@gmail.com ◦ www.josuecaraballo.com ◦ 972.415.1132

Objective

I strongly desire to build & deploy robust intelligent systems that leverage statistical & machine learning methods and tools to derive actionable business knowledge from massive, heterogenous and ever-growing datasets.

Education

Bachelor's Degree in Computer Science

University of Texas at Arlington, Arlington, TX
Spring 2013 - Fall 2017

PhD in Computer Science

University of Texas at Arlington, Arlington, TX
Fall 2019 - Spring 2025

State Farm - Data Scientist

Nov. 2019 - Present

- Began work as a TWA in Nov. 2019 - May 2020
- Exploring data to gain understanding of underlying distribution, bringing up questions to discuss with business partners.
- Work with business partners to ensure intelligent system leverages correct data elements to derive business insight.

State Farm - Hadoop Developer

Jan 2018 - Nov. 2019

- Design, development and unit/acceptance-testing of Hadoop data-processing pipelines & supporting software
- DevOps enabler, automating existing processes through automated build/test/deployment pipelines.
- Setting & meeting automated testing goals for code

www.josuecaraballo.com

Independent Web Developer

Jan 2017 - Jan 2018

- Sale of web applications to non-technical folks
- Gathering requirements and developing web-application code
- Leveraging Git during development for customer audit trail

University of Texas at Arlington

Undergraduate Researcher

May 2016 to Jan 2018

- Designed, developed and demo'd web-applications for faculty/staff/students
- Participated in various hack-a-thons/programming competitions while representing the university
- Successfully set clear requirements for software dependencies developed by others

– Mac – Windows – Linux – Docker – Kubernetes – Hadoop
– Spark – Kafka – Flume – Hive – Impala – Cloudera Manager
– Hue – LaTeX – HTML/HTML5 – CSS/CSS3 – SQL – HQL
– Java – Groovy – Scala – Clojure – Leiningen – Maven – Spring
Boot – Javascript – C/C++ – ARM Assembly – R – Python
– Jupyter – Pandas – Flask – Ruby – Perl – Bash Scripting – Regular Expressions

Experience

Software & Languages

Publications

ClaimBuster: The First-ever End-to-end Fact-checking System

ClaimPortal: Integrated Monitoring, Searching, Checking, and Analytics of Factual Claims on Twitter

Modeling Factual Claims by Frames