Nelson Abdiel Colón Vargas, Ph.D. (Nel Abdiel)

Washington,
D.C.,
20011, USA

+1 (202) 446-6256
nel.abdiel@gmail.com
http://nel.world

EDUCATION

The Data Incubator, Washington, D.C.
Certificate in Data Science — 2016
• *Accepted as a Fellow. The highly competitive program maintains a 3% acceptance rate.*

University of Iowa, Iowa City, Iowa
Ph.D. in Mathematics — 2016
M.Sc. in Mathematics — 2014

University of Northern Iowa, Cedar Falls, Iowa
M.A. in Mathematics — 2010

University of Puerto Rico, Rio Piedras Campus, San Juan, Puerto Rico
B.Sc. in Mathematics — 2009

EXPERIENCE

White House Presidential Innovation Fellows, Washington, DC.
Presidential Innovation Fellow - US Department of Veterans Affairs — Jan 2019–Present

- Presidential Innovation Fellow detailed to the Office of the CTO at the Department of Veterans Affairs leading efforts on improvement of Data Quality, Automation and Fraud Detection to better serve our Veterans.

Microsoft, Redmond, WA / Humacao, P.R.
Data & Applied Scientist II - Intellectual Property Protection & Solutions — Feb 2018–Jan 2019
Responsibilities:

- Meet weekly with Forensic Investigators and Risk Managers to discuss new trends seen in attack vectors and the thought process that lead to the detecting them.
- Turn the information gathered in meetings into quantifiable features and test the validity of said features and the detection logic.
- Build the Artificial Intelligence portion of automated solutions and present said solution to both the Development and Analytics teams.
- Present KPI measurements of newly deployed automated solutions to stakeholders.

Projects:

- Develop algorithms to detect suspicious activations based on time and geospatial data using Topological Data Analysis and Time Series Anomaly Detection techniques. Built on Python with Pandas, NumPy and SciKit-Learn among other libraries.
- Build and maintain Machine Learning models to detect malicious and scripted registrations with False Positive and False Negative rates of less than 2.5%, reducing manual revision by 87%.
- Develop SQL string matching algorithms to efficiently query entries in a database by similarity in fields of unstructured text, for example, columns corresponding to residential addresses.
- Perform statistical analysis to determine correlation among different vector attacks.
- Created Time Series models in Python to detect and prevent scripted attacks.
- Designed and built data pipelines with the use of Python and Azure to automate the ETL and analysis of data related to newly discovered fraudulent accounts.
- Automated daily data analysis revision reducing investigator's time from 30-45mins to under a minute.

Metric Geometry & Gerrymandering Group, Boston, MA
Volunteer Data Developer — Feb 2018– Sep 2018

- Contribute to the development of open source data visualization tools for geo-spatial data.
- Contribute to the development of open source tools and algorithms for determining if a district have been Gerrymandered.
- Data Science Advisor at the Voting Rights Data Institute summer program (July, 09-13, 2018), run by MGGG. Helped students turn their projects into production ready solutions to be deployed in the cloud.

Nelson Abdiel Colón Vargas, Ph.D. (Nel Abdiel)

Canóvanas, Puerto Rico, 00729, USA
+1 (787) 550-8328
nel.abdiel@gmail.com
http://nel.world

EXPERIENCE

Soteria - Security Consulting & Data Analytics, Charleston, S.C.
Data Scientist — Jun 2016–Jan 2018

Data Science and Analytics Projects:

- Developed predictive models to detect Phishing emails and websites based on content and attributes using Python packages such as Scikit-Learn, NLTK and Pandas.
- Built algorithms to detect large scale phishing campaigns with the use of clustering techniques and Natural Language Processing.
- Concocted machine learning classifiers from Scikit-Learn to detect and correctly identify different phishing techniques such as Typosquatting, Spoofing, Whaling and Spear Phishing among others.
- Created ranking system algorithms to prioritize human revision of current threats.
- Developed text classification models to assess risk of brand infringement with the use of Python packages like NLTK and WordNinja.
- Wrote MapReduce jobs to detect patterns in daily and historical domain name registrations.
- Used statistical analysis on usability data to better understand and enhance user engagement and user story.
- Develop predictive model to automate the detection of malicious websites with a 97% accuracy.
- Develop probabilistic model to predict the outcome of ping pong matches in the office.

Python Development Projects:

- Designed auto generated, data-oriented marketing material for social media with the use of Python, LaTeX and the Twitter API.
- Engineered and maintain analytics API with the used of Cassandra and Python packages like Pandas, Flask.
- Developed pipelines to autogenerate reports using Bash, Python and LaTeX.
- Developed model for detecting anomalous behavior based on user's historical data with the use of FFT.
- Engineered in-house SaaS analytics solutions to detect Spear Phishing, Whaling and other types of phishing.
- Built dashboards for internal use using Python, Flask, Django, HTML5/CSS3, Bootstrap, Bokeh and D3.js for analysts to interact with data from OSSEC and other services in a more visual matter.
- Created software for feature extraction from unstructured data from $+30000$ websites a day.
- Devised Python pipeline to automate the classification of $+30000$ websites received per day.
- Involved in the decision making and engineering of technologies and infrastructure for ETL.
- Contributed to the analysis and writing in forensic investigation reports.
- Ameliorated solutions by redesigning ETL and data analytics to function as a Linux service.
- Installed and configured OSSEC on servers, and designed and implemented new OSSEC rules to improve efficiency in data collection.

K2 Data Science, New York City, N.Y.
Data Science Curriculum Advisor — Dec 2016–Jul 2017

- Designed and developed assignments with detailed solutions for the following subjects: Data Cleaning and Preprocessing, Exploratory Data Analysis, Regression, Classification.
- Helped plan and develop comprehensive learning material to aid students with assignments and prepare them for their future roles as Data Scientists.

PUBLICATIONS

The Localized Skein Algebra is Frobenius — Oct 2017
N. Abdiel and C. Frohman, *Algebraic and Geometric Topology*, Volume 17, Issue 6.

Frobenius Algebras Derived From The Kauffman Bracket Skein Algebra — Jan 2016
N. Abdiel and C. Frohman, *Journal of Knot Theory And Its Ramifications*, Volume 25, Issue 4.