Rukmal Weerawarana

rukmal.weerawarana@gmail.com • https://rukmal.me • (206) 839-6891

Education

Master of Science (Financial Engineering) — **Hoboken, NJ**
Stevens Institute of Technology — May '19
School of Business (Division of Financial Engineering)
GPA (4.0 Scale): 3.8
Awards: Provost's Master's Fellowship

Bachelor of Arts in Business Administration (Finance) — **Seattle, WA**
University of Washington — Jun '17
Michael G. Foster School of Business (Department of Finance and Business Economics)

Certificate in Quantitative Fundamentals of Computational Finance — **Seattle, WA**
University of Washington — Aug '16
College of Arts and Sciences (Department of Applied Mathematics)

Work Experience

Software Engineer (Data Science) — **Seattle, WA**
ExtraHop Networks — Sep '19 - Present
Unusual Behaviors Group; Data Science R&D

- Designed and developed high throughput data pipelines to transport and load semi structured data across numerous globally distributed data centers. Experienced in creating enriched data sets to deliver insights to internal and external stakeholders.
- Produced numerous Data Science applications for both research and production use, with modern AWS serverless technologies and architectures. Applied modern Data Science algorithms and methodologies to Big Data to deliver insights to our customers.
- Formalized, engineered, and managed the core internal ETL pipeline for over 2 years, delivering value to numerous internal teams across the business. Scaled the core pipeline to handle an order of magnitude increase in data volume, while increasing functionality, due to increased adoption of ExtraHop's Cybersecurity product, Reveal(X) 360.
- Enhanced ExtraHop's core threat hunting ability as a detector writer, applying unsupervised learning algorithms to Big Data scale traffic flows in the cloud. Implemented numerous real-time detectors that act as early warning signals of potential bad actors on corporate networks across our customer base.
- Participated in recruitment efforts, and interviewed teammates before, and during the COVID pandemic. Assisted with mentoring and on-boarding new team members.
- ExtraHop secures hybrid cloud enterprises by providing a real-time stream processing sensor (both at the data center, and in the cloud), that transforms unstructured network packets to structured wire data at line rate. Data Science ideates, architects, deploys, and delivers a SaaS NDR solution to help secure the networks of our customers around the world.
- ExtraHop Networks was acquired by Bain Capital and Crosspoint Capital Partners for $900 million in July 2021.

Research Assistant **Hoboken, NJ**
Stevens Institute of Technology Aug '18 - May '19
Sensorimotor Control Laboratory; Department of Biomedical Engineering & Stevens Institute for Artificial Intelligence

- Designed and implemented algorithms to assess and classify tremor severity in patients with late-stage Parkinson's Disease.
- Created a highly scalable and extensible web application to be used by the researchers in the lab during this project. This web application incorporated HIPAA-compliant data storage and access, as well as efficient cluster management with Docker and Kubernetes.
- Provided input to hardware research based on statistical tremor analysis results, with the goal of designing a complete low-cost system for tremor analysis. Contributed to additive manufacturing models and designing sensor arrays for a data collection glove to be used for the tremor analysis project.
- Performed cost comparisons to identify the most economical Infrastructure-as-a-Service (IaaS) provider, given the storage, computation, and cost constraints of the project.

Summer Research Fellow **Troy, NY**
RPI-IBM HEALS Research Center May '18 - Aug '18
Tetherless World Constellation; Rensselaer Polytechnic Institute
AI Horizons Network; IBM Research

- Led the design and development of the PaperRank Framework, a methodology for deriving probabilistic community trust in academic publications. PaperRank utilized the PageRank algorithm, coupled with a Gamma Mixture Model applied to citation networks of academic publications. A proof-of-concept was implemented, from extraction to final trust score computation, analyzing over 14 Million articles from the NCBI PubMed Database.
- Formulated and implemented novel strategies for semantically-enhanced automated extraction of medical directives from Clinical Practice Guidelines (CPGs), for eventual inclusion in a knowledge graph of Diabetes diagnosis and treatment directives. Built the 'Guideline Explorer', a tool for efficiently visualizing and examining the American Diabetes Association's 2018 CPGs.
- Explored the field of 'Semantalytics', which lies at the intersection of Semantics and Analytics. Drafted a Vision statement for the future exploration of this novel field of research, through the lens of bioinformatics.
- Formulated an Electronic Health Record (EHR) simulation engine, which utilized Monte Carlo simulations based on a generalized population heuristic to vary idiosyncratic patient attributes. The EHR Simulation engine would suggest medical tests that would be statistically likely to identify previously unknown medical issues.
- Developed the 'Guideline Analysis Framework', a mathematical formulation of CPGs. This framework was designed to enable the comparison of various CPGs from differing medical authorities addressing the same set of diseases, and to detect disparities in treatment directives.

Laboratory Assistant **Hoboken, NJ**
Stevens Institute of Technology Sep '17 - Dec '18
Hanlon Financial Systems Laboratory; School of Business & Stevens Institute for Artificial Intelligence

- Spearheaded an effort to discover and implement new processes, adopt more adaptable technology, and increase functional collaboration, to help realize the teaching and research goals of the lab.
- Assisted in student project guidance, and extra-curricular instruction. Introduced interactive programming technology to aid in the Introduction to C++ course taught to MSFE students.
- Facilitated the daily operations of the lab, including assisting instructors and students (Graduate and Undergraduate), and maintaining hardware and software resources.

Business Management Team Lead **Seattle, WA**
University of Washington May '16 - Aug '17
UW Hyperloop; College of Engineering
Other Titles: Impact Development Team Lead, Control Systems Team Member

- Represented the University of Washington at the inaugural SpaceX Hyperloop Pod Competition in Hawthorne, CA. We placed 4th in the United States, and 6th globally; the competition initially received 1,700 team proposals, which were narrowed down to 30 finalists.
- Led the Business Management Team to launch a highly successful crowdfunding campaign, raising over $20,000 in cash (with an initial goal of $10,000), and over $80,000 of source materials used in the construction of the Pod. The collective effort of the team led us to have the lowest-cost Pod among the 30 final teams.
- Spearheaded the sourcing and delivery of over $50,000 of raw material, including high-density Carbon Fiber, release agents, and powerful Neodymium magnets for the final Pod assembly.
- Led sponsor outreach, and maintained relationships with internal (at the University of Washington) and external supervisors and supporters.
- Assisted the Controls Team Lead in the final design and implementation of electronic mapping and wiring on the Pod.
- Contributed to late-stage troubleshooting efforts that led to the successful implementation and deployment of Halbach arrays on the Pod, which facilitated levitation and magnetic propulsion on the test track.
- Explored the transformative economic and social effects a hypothetical Hyperloop system could have on the Pacific Northwest of the United States.

Undergraduate Research Assistant **Seattle, WA**
University of Washington Mar '15 - Jul '15
UW CubeSAT Team; Advanced Propulsion Laboratory
Other Titles: Avionics Team Lead

- Led the avionics team to design a communications software architecture paradigm for the on-board computer systems of the (previously; changed to LEO by NASA) Lunar-orbit CubeSAT.
- Investigated the effect of radiation beyond the Van-Allen belts (experienced during translunar flight). Recommended ideal processor architectures and other redundant measures that may be taken to mitigate the effect of this radiation on the flight systems of the CubeSAT.
- Proposed a system of Hamming codes to increase data transfer fidelity during data dumps from the CubeSAT to the ground station at the University of Washington. Identified specific data encoding paradigms to increase data throughput to the Earth downlink from the CubeSAT.
- Presented a final communications software architecture proposal to a panel of researchers from the Advanced Propulsion Laboratory and NASA.
- The UW CubeSAT was deployed to low-Earth orbit as secondary payload on a Cygnus cargo spacecraft, which launched aboard a Northup Grumman Antares rocket in 2019.

Software Engineering Team Lead **Seattle, WA**
ZocialGPA, Inc. Feb '15 - Jan '16
Other Titles: Software Engineering Intern

- Designed a highly scalable and efficient software ETL stack, building datasets from various social networking platforms, including Facebook, Twitter, and LinkedIn.
- Developed natural language processing and sentiment analysis algorithms to derive "social GPA" scores from a user's social profiles.
- Refactored and modularized the entire company codebase, to enable efficient component-based auto scaling with Apache Stratos and Amazon AWS.
- Implemented a mobile-first web end user interface for the platform, and redesigned the internal company management console.

Software Engineering Intern **Colombo, Sri Lanka**
WSO2, Inc. Jun '14 - Sep '14
Apache Stratos Team;

- Investigated the viability of alternate hypervisor stacks for eventual integration with the Apache Stratos Platform-as-a-Service (PaaS) framework.
- Developed a new user interface for the Stratos Manager Console using the JaggeryJS MVC framework, which was packaged and shipped with Stratos version 4.1.0.
- Conducted isolated integration tests with the CoreOS+Docker (LXC-based) hypervisor stack as an alternative to the existing hypervisor (KVM) used in Stratos.

Undergraduate Research Assistant **Seattle, WA**
University of Washington Apr '14 - Aug '14
Mullins Molecular Retrovirology Laboratory; Department of Microbiology

- Assisted the Lead Scientific Programmer of the lab in identifying and extracting sequences of mutated genes in human genome sequences.
- Developed algorithms which aided in the creation of targeted retroviral therapies for patients with HIV/AIDS.
- Created efficient and scalable data traversal algorithms for the ingestion and transformation of large human genome sequences.
- Utilized Hyperfreq, a tool for Bayesian analysis of APOBEC3G-induced hypermutations, to identify sequences of the genome likely to have been mutated.

Publications

Learned Sectors: A fundamentals-driven sector reclassification project **2019**
Rukmal Weerawarana, Yiyi Zhu, Yuzhen He
arXiv preprint; arXiv:1906.03935
https://arxiv.org/abs/1906.03935

- Market sectors play a key role in enabling the efficient flow of capital through the modern Global economy. An analysis of existing sectorization heuristics show that they are not entirely quantitatively driven, but rather are highly subjective and rooted in dogma. To this end, we introduce a new fundamentals-driven Learned Sectors heuristic.
- Using the HCA heuristic, we generate a set of 60 potential candidate learned sector universes. We then introduce reIndexer, a backtest-driven sector universe evaluation research tool, to rank the candidate sector universes produced by our learned sector classification heuristic.
- This rank was utilized to identify the risk-adjusted return optimal learned sector universe as being the universe generated under CLINK (i.e. complete linkage), with 17 sectors. The optimal learned sector universe was tested against the benchmark GICS classification universe with reIndexer, outperforming on both absolute portfolio value, and risk-adjusted return over the backtest period.

(Draft) **Inferring Community Trust from Citation Graphs** **2019**
Jamie McCusker, Rukmal Weerawarana, Alexander New, Kristin P. Bennett, Deborah L. McGuinness
https://drive.google.com/open?id=1SlSfZrwOQYPOmrKbrGLjjAFzWCk-qlwm

- We introduce the PaperRank scoring algorithm; a proxy of scientific community trust in a given publication. This score is derived from the classic PageRank algorithm (applied to academic citation networks), in conjunction with a one-dimensional Gamma Mixture Model to normalize the PageRank scores on a 3-group publication notoriety heuristic.
- The key contributions of this publication are the PaperRank Framework (currently configured for use with the NCBI PubMed Database), and a generalized algorithm for optimizing a Gamma Mixture Model (GMM).
- A key application of PaperRank would be to underwrite provenance-enabled Knowledge Graphs (where assertions are justified by academic publication references) with a 'trust' heuristic, to enable quasi-probabilistic behavior during inference and traversal.

Semantic Modeling of Cohort Descriptions in Research Studies **2018**
Shruthi Chari, Rukmal Weerawarana, Oshani Seneviratne, Jamie McCusker, Deborah L. McGuinness, Amar Das
Knowledge Representation and Semantics Workshop; AMIA 2018 Annual Symposium
https://tw.rpi.edu/web/doc/semantic_modeling_of_cohort

- This research addresses a key challenge faced by physicians using Clinical Practice Guideline recommendations; determining how well idiosyncratic cohort evidence generalizes to the greater clinical population.
- The end goal of this system is to enable the parsing of publications broadly identified as research studies to extract cohort variables and exposure/intervention groups defined within the structured population descriptions to better inform treatment decisions.

What is a Knowledge Graph? **2018**
Jamie McCusker, John S. Erickson, Katherine Chastain, Sabbir Rashid, Rukmal Weerawarana, Marcello Bax, Deborah L. McGuinness
https://drive.google.com/open?id=19uND_fkRTd_m-i-SYBznq1wuCq1IvHhO

- This work attempts to synthesize a clear and unambiguous definition of a 'Knowledge Graph' that conforms to current knowledge graph research, while constraining the research space that may be considered a knowledge graph.
- We evaluate a wide variety of knowledge resources, graphs, and other ontologies to determine if they qualify under our definition, and built a 'Knowledge Graph Catalog' to support this effort.

Talks

Neural Ordinary Differential Equations **2020**
https://drive.google.com/file/d/1fqVH6GJe1TcRyD6tL2cDUXkkkhyRq4vY
A literature review of Neural Ordinary Differential Equations by Chen et al., a new family of deep neural network models that parameterizes the hidden state of a neural network.

Learned Sectors **2019**
Collaborators: Yiyi Zhu, Yuzhen He
https://drive.google.com/file/d/12DcqxSJAR-PTLZcRqZ6SAapKzQFBuwA4
Learned Sectors project overview, covering the trained hierarchical clustering model, the reIndexer validation system, and a discussion of the final risk-adjusted return optimal sector universe.

Knowledge Graph Fundamentals **2018**
https://drive.google.com/file/d/1f21S_QZtm6aYiYLnXPA5opgsoBdIX7zX
An overview of the fundamental technology powering modern knowledge graphs, focusing on the concepts of semantic data, ontologies, and inference.

High Frequency Trading (HFT) - A Deep Dive **2017**
https://drive.google.com/file/d/1I7JuZhVzsAT84xe88lLsbHbSIa-Ev7eF
A deep dive into High Frequency Trading (HFT), covering market microstructure, exchange dynamics, regulatory implications, electronic order execution models, RegNMS, algorithmic trading, and popular HFT-driven strategies for exploiting arbitrage opportunities.

Leap into the Future with Leap Motion **2014**
http://uwhackers.github.io/leap-motion-slides/
An introduction and overview of the Leap Motion; the underlying technology, JavaScript API, and a WebSocket game demo.

Selected Knowledge Areas

Artificial Intelligence and Data Science
Anomaly Detection, Bibliometrics, Cluster Analysis, Data Visualization, Dimensionality Reduction, Graph Analytics, Knowledge Representation, Machine Learning, Natural Language Processing, Predictive Modeling, Semantic Analysis, Semi-Supervised Learning, Statistical Classification, Time Series Analysis, Unsupervised Learning

Computer Science and Software Engineering
Algorithms, Cloud Computing, Data Structures, Distributed Systems, Object Oriented Design, Operating Systems, Rapid Prototyping, Scalable System Architecture and Design, Scientific Computing, Unit Testing

Economics and Econometrics
Computational Econometrics, Credit Risk Modeling, Macroeconomics, Managerial Economics, Microeconomics, Yield Curve Modeling

Finance
Advanced Derivatives, Asset-Backed Securities, Banking and Financial Systems, Capital Budgeting, Computational Finance, Corporate Finance, Exotic Derivative Pricing, Financial Reporting, Fixed Income, Foreign Exchange Risk, International Finance, Managerial Accounting, Market Microstructure, Modern Portfolio Theory, Risk Analytics

Mathematics
Advanced Probability Theory, Combinatorics, Convex Optimization, Differential Equations, Graph Theory, Linear Algebra, Mathematical Logic, Multivariable Calculus, Non-Convex Optimization, Nonparametric Statistics, Partial Differential Equations, Real Analysis, Statistical Modeling, Stochastic Calculus

Other
AI Ethics, Elementary Computational Genomics, Elementary Quantum Computing, Sociology of Science

Skills and Technologies

Application Software
Adobe Creative Cloud, Autodesk AutoCAD, Blender, Bloomberg Terminal, Microsoft Office, Protege, SolidWorks

Databases
Apache Jena, Blazegraph, MongoDB, MySQL, Neo4j, Redis

Deployment, Orchestration, and Continuous Integration Tools
Amazon Web Services, Docker, GNU Make, Google Cloud Platform, Heroku, Kubernetes, Microsoft Azure, Travis CI

Frameworks and Libraries
Electron, Flask, Go Revel, NLTK, Node.js, R Shiny, RDFLib, SciPy, Socket.IO, Spring, TensorFlow, scikit-learn

Operating Systems
Linux (Ubuntu, Fedora, etc.), Windows, macOS

Programming and Scripting Languages
Bash, CSS, Go, HTML, Java, JavaScript, Jinja, LaTeX, MATLAB, Perl, Python, R, SPARQL, SQL

Reproducible Research Tools
GitHub, Google Colaboratory, Knitr, Microsoft Azure Notebooks, Overleaf, Project Jupyter, Read the Docs

Soft Skills
Conflict Resolution, Excellent Communication Skills, Excellent Writing Skills, Extensive Leadership Experience, Project Management, Public Speaking

Selected Projects

Precis (Jinja, LaTeX, Python, SPARQL) **2021**
https://precis.rukmal.me
Precis is an Ontology for modeling personal professional metadata. The extended Precis toolkit also includes a Pythonic search API for the Ontology, a JSON data loader, and an extensible templating engine.

fe621 (LaTeX, Python, scikit-learn) **2019**
https://git.rukmal.me/FE-621-Homework
fe621 is a Python library that provides functionality for lattice based derivative pricing models, exotic option picing, Monte Carlo simulations, numerical differentiation and integration, and optimization.

reIndexer (Python, SciPy) **2019**
https://git.rukmal.me/reIndexer
reIndexer is a research tool for the backtest-driven evaluation of different sectorization heuristics, using a system of synthetic ETFs, and efficient portfolios of those synthetic ETFs.

HTKG and NYPD-Compstat-LD (Kubernetes, Neo4j, Python) **2019**
Collaborators: Ryan Hartman, Ayush Kalla, Kovid Shukla, Sanket Saharkar
Awards: First Place [Technology] (IBM Bluehack Against Human Trafficking)
https://git.rukmal.me/NYPD-Compstat-LD
HTKG (Human Trafficking Knolwedge Graph), and NYPD-Compstat-LD (the main data ingestion engine) is a knowledge graph platform for linking suspected human trafficking advertisements with crime data from the NYPD to retroactively assess trends.

PaperRank Framework (Kubernetes, Python, Redis, SciPy, scikit-learn) **2018**
https://git.rukmal.me/PaperRank
The PaperRank framework is designed to enable bibliometrics and citation analysis of academic literature graphs. It is highly extensible, and designed to be corpus-agnostic; currently, it is configured for use with the NCBI PubMed database.

Derivative Visualizer (Knitr, R) **2017**
https://git.rukmal.me/DerivativeVisualizer
An R library for plotting simple profit/loss graphs for equity and derivative positions.

401(k) Portfolio Optimization Analysis (Knitr, LaTeX, R) **2016**
https://git.rukmal.me/UW-CFRM-462-Portfolio-Optimization
An analysis of efficient portfolios of exchange traded funds, including time series modeling of returns with Monte Carlo simulations, and heteroskedasticity analysis.

ZocialGPA Stack (Amazon Web Services, Docker, JavaScript, MongoDB, NLTK, Node.js, Python, Redis) **2015**
https://github.com/zocialgpa
ZocialGPA was a multiplatform social analytics product, which used Facebook, Twitter, and LinkedIn feeds to compute a ZGPA score for users, based on a collection of statistical NLP heuristics.

Awards and Certifications

Provost's Master's Fellowship **2017**
https://www.stevens.edu/admissions/tuition-financial-aid/graduate-funding-aid/assistantships-fellowships
Awarded to first-year Graduate students entering the Stevens Institute of Technology for graduate study.

International Baccalaureate Diploma (The British School in Colombo) **2013**
https://www.ibo.org/programmes/diploma-programme/
Higher Level: Mathematics, Physics, Chemistry; Standard Level: Geography, English, Spanish; Extended Essay: Electromagnetism (Physics).

Bloomberg Market Concepts Certificate **2017**
https://www.bloomberg.com/professional/product/bloomberg-market-concepts/
A Bloomberg Terminal certification through the lens of Economic Indicators, Currencies, Fixed Income, and Equities.

Activities

Clubs and Organizations
Delta Upsilon Fraternity, Foster Finance Association, Husky Traders, Stevens Society of Financial Engineers (SSFE), UW Financial Engineering Club (UWFEC), UW Hackers, UW Sri Lanka Student Organization, Washington Yacht Club

Hackathons
Dubhacks, Facebook Seattle Regional Hackathon, IBM Bluehack Against Human Trafficking, StartupUW Startup Weekend

Recreational Sports
Basketball, Running, Strength Training, Swimming

Volunteering Activities
Conscious Crew Volunteer, University District Food Bank Volunteer

Contact Information

6421 139th Pl NE Apt 50
Redmond, WA
98052-4588
USA

Phone: (206) 839-6891
Email: rukmal.weerawarana@gmail.com
Homepage: https://rukmal.me
LinkedIn: https://linkedin.com/in/rukmalw