# Tek Cub

## FAMILIAR TECHNOLOGIES / CONCEPTS

**Languages:** C/C++, Python, R, TeX/LaTeX, Regex, SQL
**Tooling:** DataSpell, JupyterLab, OverLeaf, Visual Studio, CMake, Catch2, Bash, Docker, Git
**Operating Systems:** Linux/Ubuntu, Windows
**Platforms:** Microsoft Azure, GitHub, Wolfram Mathematica
**Methods:** Supervised/Unsupervised Learning, Software Development Life Cycle SDLC , Test-Driven Development TDD

## DATA SCIENCE PROJECT

**Natural Language Processing (NLP)** January 2022

https://github.com/tek-cub/nlp-job-postings
(In the repository, a pseudonymous version of my resume can be found at the path *data/Resume_Tek-Cub.pdf*.)
Implemented the Manning liveProject: "Decoding Data Science Job Postings to Improve Your Resume" and completed its objectives with Python through Jupyter Notebook. See the README on GitHub for more details.
An overview of the project's scope and tasks:

- package and environment management with Anaconda
- processing data with Pandas; NLP and machine learning with Scikit-learn; use of Matplotlib and NumPy
- statistics to better understand problems, make decisions, and interpret results
- parallelism and Azure's cloud platform to reduce the main algorithm's run time by approximately 73%
- extract, transform, load the job postings dataset from web scraped HTML files with Beautiful Soup
- vectorizing text (a user's resume and job postings), by means of the TF-IDF word embedding
- ranking the similarity (cosine) of a user's resume document with entries in the dataset
- feature extraction of select data using a singular value decomposition: decreasing its dimensionality
- k-means clustering of extracted features: allowing the analysis of different types of job requirements
- visualizing contents of clusters, that is, most frequent words: to represent desired items from employers
- ranking the similarity of a user's resume to a cluster

## WORK EXPERIENCE

**Back-End Web Developer Intern** May 2017 – December 2017
Gnowit Inc. (real-time media monitoring and intelligence), Ottawa ON

- Developed server-side Java software for a distributed web application as instructed by my supervisor, highlights:
  - carried out work assignments and met their requirements by continuously learning and applying new: technologies, programming methods (concurrency and recursion), design patterns, and software systems
  - worked with application code that depended on Redis (NoSQL database) and RabbitMQ (message broker)
  - improved the maintainability and extensibility of an existing database module by refactoring its interface/API
  - created documentation for refactored software, that also described why it was refactored from its old interface: helping current and future personnel understand that part of the code base
  - automated many search and replace tasks using regular expressions, that is, after making breaking changes in my development environment, in order to reflect those changes
  - checked programs worked as intended; used debugger to resolve issues and discover how code functioned
  - tracked file changes with version-control software (Subversion) and pushed the changes I made upstream
- I created a science-fair presentation for an external audience, alongside a team of coworkers who provided feedback and multimedia content. This presentation outlined the activities of the company's interns in simple language and I engaged any interested observers/stakeholders at an event booth.

## EDUCATION

**Bachelor of Science, major in Applied Mathematics** 2025: Expected Graduation Year
Athabasca University, Athabasca AB

- courses in calculus, differential equations, discrete math, linear algebra, and statistics
- computer science electives, including data structures & algorithms