

Nicholas Attila Kovacs

DATA SCIENTIST & BIOINFORMATICIST

☎ (248) 895-2704 | ✉ Nicholas.Kovacs@gatech.edu | 🏠 NicholasAKovacs.com | 📱 NicholasAKovacs | 📺 NicholasAKovacs

Expertise

- Computational biologist with a data-driven dissertation
 - Preferred language is Python, but excited to branch out into different languages such as Julia or Java
 - Favorite packages are Pandas, Scipy, Scikit-learn, Plotly, Matplotlib, and Seaborn
 - Desires to apply good software engineering practices to solving biological problems using statistics
- Enjoys working in multidisciplinary teams and very comfortable presenting to diverse audiences
 - 5 oral and 2 poster presentations at international and domestic scientific conferences
 - Managed 8 senior undergraduates per semester as a teaching assistant for a biophysical chemistry lab course
 - 2 first-author and 6 co-author publications with computational and experimental scientists. 3rd first-author manuscript in preparation

Education

Ph.D. Bioinformatics GPA: 3.5

GEORGIA INSTITUTE OF TECHNOLOGY

Atlanta, GA

Aug 2013 - Dec 2018 (expected)

B.S. Biochemistry and Molecular Biology/Biotechnology GPA: 3.3

MICHIGAN STATE UNIVERSITY

East Lansing, MI

Aug 2008 - May 2012

Experience

Graduate Research Assistant

GEORGIA INSTITUTE OF TECHNOLOGY

Atlanta, GA

Aug 2013 - Current

Dissertation: *Data Mining the Atomic Structure of the Ribosome to Unravel the History of Protein Folding*

- **Summary:** Analyzed atomic structures of ribosomes comprised of over 400,000 atoms, using structural bioinformatics and machine learning.
- **Results:** Two 1st author research articles published. 3rd in preparation.
- **Collaboration:** Coauthored 2 experimental and 4 computational research articles.
- **Communication:** Independently wrote and awarded NSF grant to support summer research in Taiwan. Oral and poster presentations at 7 domestic and international scientific conferences.
- **Mentoring:** Awarded conference and travel funding for mentoring undergraduate student.

Atomic Interaction Network Analysis of the Ribosome

COURSE PROJECT FOR CS 7280 - NETWORK SCIENCE

Atlanta, GA

Fall 2017

- Collaborated with a team member to apply course concepts and algorithms to 3 interaction networks of the ribosome, each composed of more than 100,000 non-covalent interactions (edges) between approx. 50,000 amino acid residues (nodes).
- **Results:** Predicted RNA and protein folding domains within the ribosome by applying community detection algorithms.

Analysis and Interpretation of NGS Data from CDC

COURSE PROJECT FOR BIOL 8803B - PROGRAMMING FOR BIOINFORMATICS AND BIOL 7210 - COMPUTATIONAL GENOMICS

Atlanta, GA

Fall 2014 and Spring 2015

- Worked in multidisciplinary teams of biologists and computer scientists to analyze NGS reads provided by the CDC.
- Analyzed 97 genomic single-end and paired-end reads of *Neisseria meningitidis*, *Haemophilus influenza*, and *Haemophilus haemolyticus* generated from GAII or Illumina HiSeq/MiSeq instruments.
- **Results:** Developed a typing-tool that identifies the organism and its serotype/serogroup from fasta file inputs and constructed a genome browser of 53 annotated genomes.

Graduate Teaching Assistant

GEORGIA INSTITUTE OF TECHNOLOGY

Atlanta, GA

9 Semesters

- **Biophysical Chemistry Lab(CHEM 4582)** - 6 semesters - Instructed ~8 undergraduate students on experimental and computational protocols.
- **Macromolecular Structure(CHEM 6572)** - 2 semesters - Directed ~25 graduate students on the use of computational modelling programs.
- **Survey of Biochemistry(CHEM 3511)** - 1 semester - Guided ~40 undergraduate students to solve homework problems in weekly recitation.

Skills

Programming

Python, Perl, Bash, SQL, R, MATLAB, Javascript, HTML, CSS

Python Packages

Numpy, Pandas, Scipy, SK-Learn, Matplotlib, Seaborn, Plotly, NetworkX, igraph, Biopython, Jupyter, PySpark

Machine Learning

Linear Regression, Logistic Regression, SVM, Decision Trees, Random Forest, KNN, K-Means, PCA, Community Detection

Structural Bioinformatics

PyMOL, Maestro, PyRosetta, VMD, NAMD, AutoDock, MODELLER

Computational Genomics

de Novo Genome Assembly, SAMtools, BCFtools, VCFtools, bwa, GATK, JBrowse

OS and Software

Ubuntu, RHEL, Windows, OSX, Amazon Web Services, Microsoft Office, Adobe Illustrator, Cytoscape, Tableau, Git