

# Burak Himmetoglu

## PROFILE

- Data scientist with a computational/theoretical physics background.
- Expertise in applied mathematics (10+ years), numerical analysis (+10 years), statistics (10+ years), scientific computing (10+ years) and machine learning (5+ years).
- Expertise in high performance and cloud computing (5+ years).
- Expertise in various programming languages and parallel/distributed data processing frameworks (5+ years).

## EXPERIENCE

### Data Scientist, Serimmune 2017 - Present

- Built computational frameworks for analyzing peptide sequence data to discover disease specific sequence patterns/motifs.
- Built and deployed machine learning and feature selection frameworks for analyzing high-dimensional datasets for diagnostic panel construction.
- Employed parallel data processing pipelines with Spark to mine large biological sequence data on Google Cloud Platform.
- Built and deployed deep learning models to classify millions of short sequencing reads.
- Collaborated with laboratory scientists to develop data products and visualizations.

### High Performance Computing Consultant, University of California Santa Barbara 2015-2017

- Collaborated with research scientists and faculty to provide solutions to computationally demanding problems using high performance and cloud computing environments.
- Trained researchers in high performance and scientific computing, taught classes in Python, R and Matlab.
- Developed the RoboBohr package, a machine learning based electronic structure prediction framework for chemical discovery ([https:// github.com/bhimmetoglu/RoboBohr](https://github.com/bhimmetoglu/RoboBohr)).
- Developed applications of deep learning in medical time-series data.

**Post doctoral researcher, University of California Santa Barbara 2013-2015**

- Conducted research in computational physics and materials science.
- Developed methods and algorithms for predicting properties of semi-conductor systems at the quantum mechanical level.

**Post doctoral researcher, University of Minnesota 2010-2013**

- Conducted research in computational physics and materials science.
- Developed algorithms for numerical simulations of quantum mechanical systems and contributed to the open source materials modeling package Quantum ESPRESSO.
- Developed methods for computer modeling of solids and molecules.

**EDUCATION**

University of Minnesota, Twin Cities, MN — Ph.D. Physics 2010

Bilkent University, Ankara, Turkey — B.S. Physics 2005

**TECHNICAL SKILLS**

**Statistics and Machine Learning:**

1. Numerous classification, regression and unsupervised learning algorithms, including state-of-the-art deep learning based models.
2. Analysis of data for biological significance, including differential gene expression analysis, multiple hypothesis testing, sequence motif discovery and biological networks.

**Programming Languages:**

1. High level prototyping/scripting: Python and R with expertise in scientific and machine learning libraries. Parallel data processing using Spark and Hadoop.
2. Low level, high performance production: C/C++ with expertise in numerical libraries including BLAS, LAPACK, FFTW and parallel programming paradigms MPI, OpenMP. Familiar with GPU programming using CUDA.

**High Performance and Cloud Computing:**

1. Working knowledge of AWS and Google cloud platforms. Expertise in deep learning and big data technologies on the cloud.
2. Expertise in high performance computing environments, network file systems, schedulers and various linux operating systems.

## SELECTED LIST OF PUBLICATIONS

Authored over 30 journal publications. Complete list available in [google scholar](#).

1. Tree based machine learning framework for predicting ground state energies of molecules, *Journal of Chemical Physics* **145**, 134101 (2016), **B. Himmetoglu**.
2. Searching for high magnetization density in bulk Fe: the new metastable Fe<sub>6</sub> phase, *Journal of Physics: Condensed Matter* **27**, 016001 (2014), K. Umemoto, **B. Himmetoglu**, J.-P. Wang, R. M. Wentzcovitch, and M. Cococcioni.
3. Hubbard corrected DFT energy functionals: LDA+U description of correlated systems, *International Journal of Quantum Chemistry* **114**, 14 (2013), **B. Himmetoglu**, A. Floris, S. de Gironcoli, and M. Cococcioni.
4. Instability of anisotropic cosmological solutions supported by vector fields, *Physical Review Letters* **102**, 111301 (2009), **B. Himmetoglu**, C. R. Contaldi, and M. Peloso.