Capstone experience in data science

DS master’s program

February 2021
The capstone experience is a requirement of the master’s program.

It is the student’s responsibility to reach out to potential supervisor for the capstone, following all the required steps in the process.
Process
1. Find and contact a research supervisor who will agree to supervise the capstone experience. The research supervisor must be a JHU faculty member.
2. Complete a proposal form, describing the project goals, and submit to their academic advisor, who will in turn send it to the Internal Oversight Committee for approval.
# Proposal Request for the Capstone Experience in Data Science (EN.553.806)

<table>
<thead>
<tr>
<th>Student’s First Name</th>
<th>Student’s Last Name</th>
<th>SIS #</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>JHU Email</th>
<th>Alternate Email</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Data Science Program Advisor’s Name</th>
<th>Capstone Experience Supervisor’s Name (if different from advisor)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Advisor’s Email</th>
<th>Supervisor’s Email</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Area of Focus</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Title of Research Proposal</th>
</tr>
</thead>
</table>

## Description of Research - provide enough details for your proposal to be adequately evaluated (approximately 200 words)

<table>
<thead>
<tr>
<th>Completion Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Student Signature</th>
<th>Date</th>
</tr>
</thead>
</table>

The proposal described above is approved. This project meets the requirements for EN.553.806

<table>
<thead>
<tr>
<th>Capstone Supervisor Approval</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic Advisor Approval</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Science Program Director Approval</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. The proposal will include the following and must be submitted using the approved proposal request form (above):
   • Title of proposed project
   • Project description, with sufficient details (e.g., 200 words)
   • Completion timeline
   • Name(s) and signature(s) of faculty supervisor(s)
4. Upon approval, register for EN.553.806: Capstone Experience in Data Science.
5. As part of the experience, the student must write a paper or research report that must be approved by the research supervisor.

The final paper should be 6-12 pages in latex full-page format (1 inch margins, 12 point Times font) or ms-word equivalent.
6. The written paper will be summarized in a poster presented in a poster session organized at the end of each semester.
Group projects

• Group projects are acceptable to fulfill the capstone experience, with the following requirements:

  • The group submits a proposal that, in addition to the aforementioned requirements, describes the composition of the group, and the role expected by each group member in the completion of the project.

  • Groups should not include more than four students

  • Each student in the group submits an individual report, summarizing the full project, but focusing primarily on the part that was under their responsibility. The final poster may be represented individually or as a group.

• The option to complete a group project is solely at the discretion of the research supervisor.
Industrial projects

• Projects may be completed in an industrial environment, provided all previously described steps are followed.

• In particular, the presence of a JHU faculty liaison is still a requirement, in addition to a possible supervisor in the industrial context.
Timeline (3 semester case)

• Start contacting advisors during the second semester in the program, with a goal to start working on the capstone at or before the beginning of the third semester.

• Dedicate most of the third semester to the capstone.

• Recall that students should not register to 553.806 until their capstone proposal has been accepted.
Some potential advisors
Potential supervisors

• The following JHU faculty members have all expressed their willingness to be contacted for a capstone experience (without commitment to provide one).

• This list is not exhaustive (any JHU faculty may act as supervisor) but will provide entry points for students who haven’t had any direct contact with faculty yet.

• It is up to the student to reach out to potential supervisors

• What to include in your email:
  • Your name
  • The courses you have taken so far and are taking in the program
  • Why you are interested in working with the person you are contacting

• YOU SHOULD NOT CONTACT MORE THAN ONE OR TWO PERSONS AT A TIME
Raman Arora

- Machine Learning: Provable methods for deep learning and representation learning, subspace learning, multiview learning, streaming algorithms for kernel methods, online learning
- Stochastic Optimization: Non-convex optimization, stochastic approximation for large-scale problems, robust adversarial learning
- Differential Privacy: Computational tradeoffs in private machine learning, local $\approx$learning, federated learning, privacy in streaming algorithms and continual release models
- https://www.cs.jhu.edu/~raman/Home.html
Amitabha Basu

• Optimization (with current emphasis on Integer Programming)
• Discrete and combinatorial geometry
• Convex analysis
• Applications to Operations Research, Astronomy and Data Science.
• http://www.ams.jhu.edu/~abasu9/
Alexis Battle

- Human genome analysis.
- https://www.bme.jhu.edu/people/faculty/alexis-battle/
Maxim Bichuch

- Financial mathematics.
- https://engineering.jhu.edu/ams/faculty/maxim-bichuch/
Vova Braverman

• Efficient algorithms and their applications to machine learning
• Systems and networks
• Cosmological simulations
• Computational medicine
• https://www.cs.jhu.edu/~vova/
Tamás Budavári

- Computational statistics
- Bayesian inference
- Low-dimensional embeddings
- Streaming algorithms
- Parallel processing on GPUs
- Scientific databases
- Survey astronomy
- https://engineering.jhu.edu/ams/faculty/tamas-budavari/
Brian Caffo

• Statistical computing
• Statistical modeling
• Computational statistics
• Multivariate and decomposition methods
• Statistics in neuroimaging and neuroscience.
• https://www.jhsph.edu/faculty/directory/profile/1010/brian-s-caffo
Adam Charles

- Data science
- Signal processing
- Machine learning
- Computational neuroscience
- Computational imaging
- Cybernetics
- https://www.bme.jhu.edu/people/faculty/adam-charles/
Nicolas Charon

• Shape analysis
• Image analysis
• Riemannian and discrete geometry
• Medical imaging
• Computational anatomy
• http://www.cis.jhu.edu/~charon/
Jason Eisner

• Natural language processing
• Combinatorial algorithms
• Probabilistic models of linguistic structure
• Declarative specification of knowledge and algorithms.
• https://www.cs.jhu.edu/~jason/
Jean Fan

• Cell biology
• Computational biology
• Machine learning for multi-omic data
• https://www.bme.jhu.edu/people/faculty/jean-fan/
Mahyar Fazlyab

- Control
- Optimization
- Machine Learning

https://engineering.jhu.edu/ece/faculty/mahyar-fazlyab/
Elana Fertig

- Systems biology and bioinformatics
- Single cell technologies and algorithm development
- Pattern detection and semi-supervised algorithms
- Precision medicine in cancer
- https://fertiglab.com
Helyette Geman

• Finance of commodities
• Insurance
• Probability theory.
• https://engineering.jhu.edu/ams/faculty/helyette-geman/
Edinah Gnang

- Discrete mathematics
- Graph theory
- Multilinear algebra
- Image analysis
- Experimental math.

http://www.ams.jhu.edu/~egnang/
Michael Kazhdan

- Computer Graphics
- Surface Reconstruction
- Image and Geometry Processing
- [https://www.cs.jhu.edu/faculty/michael-misha-kazhdan/](https://www.cs.jhu.edu/faculty/michael-misha-kazhdan/)
Yannis Kevrekidis

• Algorithms, data, and the computer-assisted modeling of complex dynamical systems
• https://engineering.jhu.edu/chembe/faculty/yannis-kevrekidis/
Sergey Kushnarev

- Applied Statistics
- Shape Analysis
- Pattern Theory
- Medical Imaging

https://engineering.jhu.edu/ams/faculty/sergey-kushnarev/
Zachary Lubberts

- Statistics
- Optimization
- Statistical Inference on Graphs
- Functional Analysis
- Wavelet Analysis

https://engineering.jhu.edu/ams/faculty/zachary-lubberts/
Mauro Maggioni

- Analysis
- Partial Differential Equations
- Algebraic Topology
- Big Data
- Data Intensive Computation
- Harmonic Analysis over Manifolds and over Discrete Structures
- https://engineering.jhu.edu/ams/faculty/mauro-maggioni/
Enrique Mallada

• Networked dynamics and distributed systems, and their applications in engineering and science.
• Control, dynamical systems and optimization.
• https://mallada.ece.jhu.edu
Mario Micheli

- Mathematical image processing
- Shape analysis
- Variational methods
- Applied Statistics
- Mathematical 3D printing
- https://engineering.jhu.edu/ams/faculty/mario-micheli/
Daniel Q. Naiman

- Statistics
- Computational Probability
- Bioinformatics
- https://engineering.jhu.edu/ams/faculty/daniel-q-naiman/
Vishal M. Patel

• Signal and image processing
• Computer vision
• Machine learning
• Biometrics
• Biomedical image analysis
• https://engineering.jhu.edu/ece/faculty/vishal-m-patel/
Carey Priebe

- Computational Statistics
- Kernel and Mixture Estimates
- Statistical Pattern Recognition
- Statistical Image Analysis
- Dimensionality Reduction
- Model Selection
- Statistical Inference for High-Dimensional and Graph Data
- https://engineering.jhu.edu/ams/faculty/carey-priebe/
Fadil Santosa

- Image Processing
- Computer-Assisted methods
- Inverse problems (Differential equations)
- MACHINE learning
- OPTIMAL designs (Statistics),
- PHOTONICS — Research, Signal processing — Digital techniques — Mathematical model
- https://engineering.jhu.edu/ams/faculty/fadil-santosa/
Ilya Shpitser

- Causal inference
- Missing data
- Graphical models
- Algorithmic fairness
- Statistical inference in semi-parametric models.

https://www.cs.jhu.edu/~ilyas/
James C. Spall

- Stochastic Systems
- Parameter Estimation
- Stochastic Optimization
- Monte Carlo Methods and Simulation
- Neural Networks
- Control Systems
- System Identification and Kalman Filtering
- Mathematical Statistics
- Optimization Theory
- Uncertainty Calculation
- https://engineering.jhu.edu/ams/faculty/james-c-spall/
Jeremias Sulam

- Inverse problems in computer vision and signal processing
- Dictionary learning
- Machine learning and deep learning and their application to biomedical problems
- https://www.bme.jhu.edu/people/faculty/jeremias-sulam/
Trac D. Tran

• Compressed sensing, sparse recovery, and sparsity-based signal processing
• Signal representation, signal decomposition, time-frequency and time-scale analysis
• Filter banks, wavelets, multirate systems, transforms
• Audio/image/video processing, compression, adaptive filtering, and DSP for communications
• https://engineering.jhu.edu/ece/faculty/tran-trac-d/
Archana Venkataraman

- Functional Neuroimaging (fMRI, EEG)
- Machine Learning & Probabilistic Inference
- Network Modeling of the Brain
- Integration of Imaging, Genetics and Behavioral Data
- https://engineering.jhu.edu/ece/faculty/archana-venkataraman/
René Vidal

- Mathematics of deep learning
- Clustering and factor analysis
- Computer vision
- Biomedical image analysis
- Dynamical systems
- Robotics
- Signal processing
- http://www.cis.jhu.edu/~rvidal/
Soledad Villar

- Optimization for data science
- Machine learning and optimization
- Representation learning and graph neural networks
- https://www.ams.jhu.edu/villar/
Joshua T. Vogelstein

• Brain connectomics
• Statistics on graphs
• Machine learning
• https://www.bme.jhu.edu/people/faculty/joshua-t-vogelstein/
Yanxun Xu

- Bayesian statistics
- Cancer genomics
- Clinical trial design
- Graphical models
- Nonparametric Bayesian statistical inference for big data analysis
- High-throughput genomic data
- Proteomics data

https://engineering.jhu.edu/ams/faculty/yanxun-xu/
Laurent Younes

• Statistical properties of graphical models
• Mathematical and computational analysis of shape spaces
• Statistical learning methods in computational biology
• https://engineering.jhu.edu/ams/faculty/laurent-younes/