UNIVERSITY REGISTRAR
VIRGINIA TECH

CM-8156 APPROVED
GCC: 4.27.23
CGPSP: 5.3.23

ENTERED 6.5.23 ~BS

Coversheet - Proposal for New and Revised Courses

(Use for non-Pathways courses)

For CLE/Pathways courses, form can be found here: https://www.pathways.prov.vt.edu/proposal-forms.html

General Information			
Proposal Date:	2/1/23	*Department:*	Computer Science
Course Designator and Number (Cross-listed Course Designator and Number):	CS 5805-5806		
Title of Course:	Machine Learning	*Credit Hours:*	3-3
Course Transcript (ADP) Title (30 Characters & Spaces Maximum):	Machine Learning		
Instructor and/or Departmental Contact:	Trey Mayo - Director of Graduate Programs		
Contact Phone:	X0780	*Contact E-mail:*	treymayo@vt.edu
Please refer to Office of University Registrar for guidelines and policy requirements: https://registrar.vt.edu/governance.html			

Please count this course toward the following Scorecard Metrics areas:

☐ Study Abroad ☐ Service Learning ☐ Experiential ☐ Undergraduate Research

Scorecard Metrics Definitions can be found here: https://registrar.vt.edu/faculty-toolbox/scorecard-metrics.html

Please insert an X if this course should count toward First Year Experience:

☐ First Year Experience (FYE) *Include approval letter from FYE Director. More information can be found here:* http://www.fye.vt.edu

Select **ONE** of the following boxes			
X	New Course		*Revised Course (*Revision > 20%____ Revision < 20%_____*)

*For **CLE/Pathways** courses, form can be found here*: https://www.pathways.prov.vt.edu/proposal-forms.html

Please include a summary of course revisions to the Justification section of proposal

A:	***Attach statement from Dean or Departmental Representative*** as to whether teaching this course will require or generate the need for additional departmental resources.			
B:	***Attach appropriate letters of support*** (e.g., prerequisite, corequisite, or cross-list memo) from affected departments and/or colleges.			
C:	***Effective Semester:***	Fall 2023		
D:	***Change in Title From:***			
	To:			
E:	***Change in Transcript Title (ADP) From:***		***To:***	
F:	***Change in Credit Hours From:***		***To:***	
G:	***Change in Lecture and/or Lab Hours From:***		***To:***	
H:	***Course Number(s) and Title(s) to be deleted from the Catalog with APPROVAL:***			

Approval Signatures			
Department Representative	Please see attached coversheet.	*Date*	
College Curriculum Committee Rep		*Date*	
College Dean or Designee		*Date*	

UNIVERSITY REGISTRAR
VIRGINIA TECH

Coversheet - Proposal for New and Revised Courses

(Use for non-Pathways courses)

For CLE/Pathways courses, form can be found here: https://www.pathways.prov.vt.edu/proposal-forms.html

General Information			
Proposal Date:	2/1/23	*Department:*	Computer Science
Course Designator and Number (Cross-listed Course Designator and Number):	CS 5805-5806		
Title of Course:	Machine Learning	*Credit Hours:*	3-3
Course Transcript (ADP) Title (30 Characters & Spaces Maximum):	Machine Learning I, Machine Learning II		
Instructor and/or Departmental Contact:	Trey Mayo - Director of Graduate Programs		
Contact Phone:	X0780	*Contact E-mail:*	treymayo@vt.edu
Please refer to Office of University Registrar for guidelines and policy requirements: https://registrar.vt.edu/governance.html			

Please count this course toward the following Scorecard Metrics areas:

☐ Study Abroad ☐ Service Learning ☐ Experiential ☐ Undergraduate Research

Scorecard Metrics Definitions can be found here: https://registrar.vt.edu/faculty-toolbox/scorecard-metrics.html

Please insert an X if this course should count toward First Year Experience:

☐ First Year Experience (FYE) *Include approval letter from FYE Director. More information can be found here:* http://www.fye.vt.edu

Select **ONE** of the following boxes			
X New Course		*Revised Course	(*Revision > 20%*_____ *Revision < 20%*_____)

*For **CLE/Pathways** courses, form can be found here*: https://www.pathways.prov.vt.edu/proposal-forms.html

Please include a summary of course revisions to the Justification section of proposal

A:	***Attach statement from Dean or Departmental Representative*** as to whether teaching this course will require or generate the need for additional departmental resources.			
B:	***Attach appropriate letters of support*** (e.g., prerequisite, corequisite, or cross-list memo) from affected departments and/or colleges.			
C:	***Effective Semester:***	Fall 2023		
D:	***Change in Title From:***			
	To:			
E:	***Change in Transcript Title (ADP) From:***		***To:***	
F:	***Change in Credit Hours From:***		***To:***	
G:	***Change in Lecture and/or Lab Hours From:***		***To:***	
H:	***Course Number(s) and Title(s) to be deleted from the Catalog with APPROVAL:***			

Approval Signatures			
Department Representative	DocuSigned by: Cliff Shaffer 5ABBF057372B402...	*Date*	3/22/2023
College Curriculum Committee Rep		*Date*	3/23/2023
College Dean or Designee		*Date*	3/23/2023

UNIVERSITY REGISTRAR
VIRGINIA TECH

Coversheet - Proposal for New and Revised Courses
(Use for non-Pathways courses)

Course Information

Catalog Description

5805:
Provides an introduction to the field of machine learning (or data mining) and explores the common tasks of machine learning to include preprocessing, classification, clustering, the discovery of association rules/sequential patterns, and anomaly detection. Introduces fundamentals of probability theory and random variables for classification and clustering. Investigates multiple linear and nonlinear regression models to classify social phenomena. Apply application of machine learning in solving real work problems. Credit will not be given for both CS/STAT 5525 and CS 5805. Pre: Graduate standing. (3H, 3C).

5806:
Provides an in-depth understanding of classical machine learning theory using Bayesian models, statistical machine learning and pattern recognition techniques, advanced machine learning methods, and their applications. Pre: CS 5805. (3H, 3C).

Learning Objectives

CS 5805:
Having successfully completed this course, the student will be able to:

1. Use various techniques related to preprocessing prior to the use of machine learning models.
2. Describe the probability theory and random variables.
3. Identify the common tasks in machine learning/data mining models for clustering.
4. Analyze multiple linear and nonlinear regression.
5. Describe the algorithms, theories, and applications related to machine learning/data mining for classification.
6. Detect anomaly/outlier behavior and the treatment techniques.

CS 5806:
Having successfully completed this course, the student will be able to:

1. Utilize mathematical tools in theoretical machine learning
2. Identify classical machine learning methods and be able to implement them from scratch
3. Formulate and implement advanced machine learning methods
4. Analyze the machine learning models in terms of generalization, performance and scalability
5. Design generative and discriminative machine learning models for real problems
6. Utilize machine learning toolboxes and platforms for model development

Justification

For 5805:
We continue to see explosive growth in the amount of data collected across all of society. This has created an unprecedented opportunity for machine learning (also referred to as data analytics or data mining), which is the process of efficiently supervised or unsupervised discovery of interesting and useful information from collections of data.

CS 5805 is taught at the 5000 level because it builds upon general programming and analytical skills found in undergraduate courses and degrees related to computer science, statistics, electrical or computer engineering, or mathematics. The course provides the foundational knowledge needed to pursue advanced graduate study in machine learning including CS 5806 - Machine Learning. CS 5805 will also provide a suitable foundation for courses in Deep Learning, Computer Vision, and Natural Language Processing.

For 5806:
Machine learning methods have enabled significant progress in real-world applications from text classification and machine translation to image and scene understanding, robotics, etc. Advanced methods such as Bayesian inference and probabilistic models have found increasing interest both in industry and academia, as they allow encoding domain knowledge, quantifying uncertainty, and sampling from the underlying data distribution. It is therefore essential to train students with the fundamental knowledge and practical skill sets in both discriminative and generative machine learning approaches. This course provides in-depth understanding of concepts, state-of-the-art techniques and algorithms, and real-world applications of advanced machine learning. Students will gain extensive exposure to both mathematical foundations and practical development of advanced machine learning methods. The course will prepare students to tackle challenging tasks and real-world problems with different types of data (text, images, graphs, etc.) and learning tasks (supervised, unsupervised, semi-supervised, weakly-supervised, self-supervised, etc.). Through this course, students will gain mathematical intuition on how each machine learning method works, form mathematical connections between different machine learning algorithms, learn how to transform practical problems into formalized learning objectives, and how to encode priors and other forms of structure into probabilistic models.

CS 5806 is taught at the 5000 level because it builds upon general programming and analytical skills found in undergraduate courses related to linear algebra, probability, statistics, algorithm design and analysis, and computer programming. It further expands upon the material presented in its [proposed] prerequisite of CS 5805 - Machine Learning.

Prerequisites and Corequisites

For CS 5805:
Pre: Graduate Standing

For CS 5806:
Pre: CS 5805 Machine Learning

Texts and Special Teaching Aids

5805 Required textbooks:

- Raschka, S. & Mirjalili, V. (2019). *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2* (3rd ed.). Packt Publishing. pp 772.
- Tan, P., Steinbach, M., Karpatne A., & Kumar, V. (2018). *Introduction to data mining* (2nd ed.). Pearson. pp. 864.

5806 Required textbooks:

- Bishop, C. M. (2006). *Pattern recognition and machine learning.* Springer. pp. 758.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press. https://www.deeplearningbook.org/
- Murphy, K. P. (2023). *Probabilistic Machine Learning: Advanced topics.* MIT Press. https://probml.github.io/pml-book/book2.html

Supplemental material for 5806:

- Karpatne, A., Kannan, R., & Kumar, V. (2022). *Knowledge-guided machine learning: Accelerating discovery using scientific knowledge and data.* CRC Press. pp. 472.

Topic Syllabus

Topic | Percent of Course

CS 5805:

Probability theory and random variables • Random variables and Probability	10%
Preprocessing • Dataset cleaning & Attribute transformation • Attribute extraction & dimension reduction • Similarities and distances	15%
Multiple linear and nonlinear regression	10%
Classification Analysis I • Logistic regression • Decision tree • KNN	15%
Classification Analysis II • Random Forest • SVM • Naïve Bayes • Neural Network	20%
Clustering/Association Analysis • K-mean • Density-based spatial clustering with noise (DBSCAN) • Gaussian mixture model algorithm • Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) • Mean-shift • Association Rule mining AIS, SETM, Apriori, FP-tree	20%
Anomaly Detection	10%
	Total: 100%

CS 5806:

Overview and Learning Theory • Maximum Likelihood (MLE) / Maximum A-Posteriori (MAP) • Bias-Variance tradeoff • PAC Learning – VC Dimension, etc.	10%
Classical Machine Learning Methods • Bayesian Linear Models for Regression and Classification • Kernel Methods and Max-Margin Classifiers • Mixture Models; Expectation Maximization • Graphical Models • Structured Predictions (Hidden Markov Models, Viterbi Algorithm, etc.) • Variational Bound and Approximate Inference	**40%** 10% 5% 5% 10% 5% 5%
Advanced Machine Learning Methods • Deep Neural Networks • Markov Decision Processes, Bandits and Reinforcement Learning • Generative Models • Different Types of Learning: Unsupervised, Semi-supervised, Weakly-supervised, Self-supervised, Distantly-supervised	**35%** 15% 5% 10% 5%
Design and implementation of Machine Learning methods by using existing commonly- used Machine Learning toolboxes and platforms	15%
	Total: 100%

Old (Current) Topic Syllabus
N/A

Department of Computer Science
620 Drillfield Drive
Torgersen Hall, Suite 3210
Blacksburg, Virginia 24061
P: (540) 231-0780
treymayo@vt.edu

March 22, 2023

To: Course Approval Committees

From: Trey Mayo, Ed.D. Trey Mayo Digitally signed by Trey Mayo Date: 2023.03.22 11:13:07 -04'00'
Director of Graduate Programs
Computer Science

The Department of Computer Science is requesting approval of a new course proposal CS 5805-6 - Machine Learning ~~I and II~~. No new resources will be required in order to offer these courses.

TM 4/14/23

COLLEGE OF ENGINEERING
COMPUTER SCIENCE
VIRGINIA TECH

Department of Computer Science
2000A Torgersen Hall
Blacksburg, Virginia 24061-0106
shaffer@vt.edu, 540-231-4354
http://people.cs.vt.edu/shaffer

February 24, 2023

To: College and University Curriculum Committees
RE: Computer Science Course Revisions Package

The Department of Computer Science presents a coordinated package of new course proposals and course revisions for the purpose of reorganizing our graduate-level offerings in the core Computer Science domain of Machine Learning and related topics. Our current offerings are the result of uncoordinated individual actions often made in conjunction with other departments over several years. The result is a collection of courses with overlaps and inefficiencies that lead to confusion for our students.

This package is centered around a two-semester sequence directly covering the core of traditional Machine Learning topics, with the first course in the sequence also serving as the prerequisite for the other, related courses. This allows us to avoid duplicating background material across these courses. We support the core with a collection of three courses (one existing, two new) that span the generally recognized major topics related to Machine Learning: Natural Language Processing, Learning-based Computer Vision, and Deep Learning. As part of the overhaul, we break some existing cross-listings to courses long recognized as duplicative.

Our package includes the following new courses and course revisions.

- Two course sequence CS 5805-6 Machine Learning. CS 5805 partially duplicates existing course CS/STAT 5525. However, we do not currently seek any changes to CS/STAT 5525 since this course is presently integrated into other certificates and programs. Instead, we have an agreement with Statistics to support their future changes to CS/STAT 5525 to bring it more in line with their needs. CS 5806 will (within our curriculum) replace the role currently held by ECE 5424/CS 5824.
- We request to break the cross-listing agreement for CS 5824, leaving this as ECE 5424.
- We include a revision to CS 5814 Introduction to Deep Learning. In addition to minor topics list changes, CS 5805 will become the prerequisite course for CS 5814.
- We request to break the cross-listing agreement for ECE 6524/CS 6524 Deep Learning. Despite the names, this is largely duplicative with CS 5814. Breaking this cross listing will reduce existing confusion for both CS and ECE students.
- A new course proposal for CS 5624 Natural Language Processing. This course will have CS 5805 as a prerequisite.
- A new course proposal for CS 5864 Learning-based Computer Vision. This course will have CS 5805 as a prerequisite.

Implementing these proposals will leave our department with a collection of courses that properly represents

the major sub-fields within the broad area of Machine Learning and Artificial Intelligence, as is typical for major Computer Science Departments across the US.

We note that while there is duplication between CS 5805 and CS/STAT 5525, and between CS 5806 and ECE 5424, this should have relatively minor impact on the teaching loads for the departments or the health of any of these courses. Historically, we have taught multiple sections of each of these courses every year, involving instructors from each of the three departments. None of these courses will lack for students for the foreseeable future.

Implementing this package will require no new resources. The NLP and Learning-based Computer Vision courses are already in our teaching rotation having been run as special topics courses. So, all of these courses have already been taught in various forms, and are already built into our course offering structure.

Unrelated to the described collection of proposals, we also offer a new course proposal for CS 5784 Software Project Management. This course has been piloted twice already.

Sincerely,

Clifford A. Shaffer
Professor and Associate Department Head
for Graduate Studies