MATH M42DS: MATHEMATICS OF MACHINE LEARNING FOR DATA SCIENCE

Originator
brendan_purdy

Co-Contributor(s)

Name(s)
Nava, Michael (michael_nava2)
Ogimachi, Tom (togimachi)

College
Moorpark College

Attach Support Documentation (as needed)
M42DS_UC Courses.pdf

Discipline (CB01A)
MATH - Mathematics

Course Number (CB01B)
M42DS

Course Title (CB02)
Mathematics of Machine Learning for Data Science

Banner/Short Title
Math Machine Learning - DS

Credit Type
Credit

Honors
No

Start Term
Spring 2022

Catalog Course Description
Introduces machine learning algorithms with linear algebra for data science. Emphasizes the mathematical foundations of ensemble methods, discriminant analysis, deep learning, and neural networks as well as the ethical use of data. Covers applications of algebraic programming for data science.

Taxonomy of Programs (TOP) Code (CB03)
1701.00 - Mathematics, General

Course Credit Status (CB04)
D (Credit - Degree Applicable)

Course Transfer Status (CB05) (select one only)
A (Transferable to both UC and CSU)

Course Basic Skills Status (CB08)
N - The Course is Not a Basic Skills Course

SAM Priority Code (CB09)
E - Non-Occupational

Course Cooperative Work Experience Education Status (CB10)

N - Is Not Part of a Cooperative Work Experience Education Program

Course Classification Status (CB11)

Y - Credit Course

Educational Assistance Class Instruction (Approved Special Class) (CB13)

N - The Course is Not an Approved Special Class

Course Prior to Transfer Level (CB21)

Y - Not Applicable

Course Noncredit Category (CB22)

Y - Credit Course

Funding Agency Category (CB23)

Y - Not Applicable (Funding Not Used)

Course Program Status (CB24)

1 - Program Applicable

General Education Status (CB25)

B - Satisfies Math/Quantitative Reasoning req (CSUGE-B B4, IGETC 2, or 4-yr)

Support Course Status (CB26)

N - Course is not a support course

Field trips

Will not be required

Grading method

(L) Letter Graded

Alternate grading methods

(O) Student Option- Letter/Pass
(P) Pass/No Pass Grading

Does this course require an instructional materials fee?

No

Repeatable for Credit

No

Is this course part of a family?

No

Units and Hours

Carnegie Unit Override

No

In-Class

Lecture
Minimum Contact/In-Class Lecture Hours
52.5
Maximum Contact/In-Class Lecture Hours
52.5

Activity
Minimum Contact/In-Class Activity Hours
0
Maximum Contact/In-Class Activity Hours
0

Laboratory
Minimum Contact/In-Class Laboratory Hours
0
Maximum Contact/In-Class Laboratory Hours
0

Total in-Class

Total in-Class
Total Minimum Contact/In-Class Hours
52.5
Total Maximum Contact/In-Class Hours
52.5

Outside-of-Class

Internship/Cooperative Work Experience

Paid
Minimum Paid Internship/Cooperative Work Experience Hours
0
Maximum Paid Internship/Cooperative Work Experience Hours
0

Unpaid
Minimum Unpaid Internship/Cooperative Work Experience Hours
0
Maximum Unpaid Internship/Cooperative Work Experience Hours
0

Total Outside-of-Class

Total Outside-of-Class
Minimum Outside-of-Class Hours
105
Maximum Outside-of-Class Hours
105

Total Student Learning

Total Student Learning
Total Minimum Student Learning Hours
157.5
Total Maximum Student Learning Hours
157.5

Minimum Units (CB07)
3
Maximum Units (CB06)
3

Prerequisites

CS M10DS and MATH M25C and MATH M31 and MATH M15 OR MATH M15H

Advisories on Recommended Preparation

MATH M21 or MATH M35

Entrance Skills

Entrance Skills

CS M10DS, MATH M15 OR MATH M15H, MATH M25C, and MATH M31

Prerequisite Course Objectives

CS M10DS-deploy basic statistical concepts that data scientists need to know like measure of central tendency, percentiles, probability distributions, dimensionality reduction, over and under sampling, and Bayesian statistics.
CS M10DS-distinguish fundamental aspects of machine learning algorithms.
CS M10DS-frame problems to enable suitable solutions via machine learning.
CS M10DS-train (process of modeling) and evaluate machine learning models.
CS M10DS-deploy machine learning models into operations.
CS M10DS-build prediction, categorization and recommendation APIs (application program interface).
CS M10DS-deploy tools for collaborative and social programming.
CS M10DS-generate high-quality graphical and textual results.
CS M10DS-find similar items, pattern discovery (Neural Network and Machine Learning), mining data streams, frequent itemsets, link analysis, and mining graph data.
CS M10DS-distinguish supervised machine learning algorithms like k nearest neighbors, decision trees, Naive Bayes, regression, and support vector machines.
CS M10DS-distinguish unsupervised machine learning (data mining) algorithms like classification, k-means clustering, evaluation of clustering, hierarchical clustering, spectral (partitional) clustering, and Neural networks.
MATH M15-summarize data graphically by displaying data using methods from descriptive statistics, interpreting data in tables graphically by using histograms, frequency distributions, box-and whisker plots (five-number summary); find measures of central tendency for data sets: mean, median, and mode; find measures of variation for data sets: standard deviation, variance, and range; determine relative positions of data and distinguish among scales of measurements and their implications; distinguish between populations and samples; and identify the standard method of obtaining data and the advantages and disadvantages of each.
MATH M15-find simple probabilities and probabilities of compound events and compute probabilities using the complement, discrete probability distributions; apply concepts of sample space, and the binomial probability distribution.
MATH M15-standardize a normally distributed random variable; use normal distribution tables to find probabilities for normally distributed random variables and the t-distribution; use the Central Limit Theorem to find probabilities for sampling distributions.
MATH M15-construct and interpret confidence intervals for proportions and means.
MATH M15-identify the basics of hypothesis testing and perform hypothesis testing for means, proportions and standard deviations from one population, and difference of means and proportions from two populations, including finding and interpreting p-value and examining Type I and Type II error.
MATH M15-find linear least-squares regression equations for appropriate data sets; graph least-square regression equations on the scatter plot for the data sets; find and apply the coefficient of correlation.
MATH M15-use the chi-square distribution to test independence and to test goodness of fit.
MATH M15-conduct a one-way Analysis of Variance (ANOVA) hypothesis test.
MATH M15-select an appropriate hypothesis test and interpret the result using p-value; use appropriate statistical technique to analyze and interpret applications based on data related to business, social sciences, psychology, life sciences, health sciences or education, and interpret results using technology-based statistical analysis.
MATH M15H-summarize data graphically by displaying data using methods from descriptive statistics, interpreting data in tables graphically by using histograms, frequency distributions, box-and whisker plots (five-number summary); find measures of central tendency for data sets: mean, median, and mode; find measures of variation for data sets: standard deviation, variance, and range; determine relative positions of data and distinguish among scales of measurements and their implications; distinguish between populations and samples; and identify the standard method of obtaining data and the advantages and disadvantages of each.
MATH M15H-find simple probabilities and probabilities of compound events and compute probabilities using the complement, discrete probability distributions; apply concepts of sample space, and the binomial probability distribution. Honors: And the Poisson distribution.
MATH M15H-standardize a normally distributed random variable; use normal distribution tables to find probabilities for normally distributed random variables and the t-distribution; use the Central Limit Theorem to find probabilities for sampling distributions.
MATH M15H-construct and interpret confidence intervals for proportions and means.
MATH M15H-identify the basics of hypothesis testing and perform hypothesis testing for means, proportions and standard deviations from one population, and difference of means and proportions from two populations, including finding and interpreting p-value and examining Type I and Type II error.
MATH M15H-find linear least-squares regression equations for appropriate data sets; graph least-square regression equations on the scatter plot for the data sets; find and apply the coefficient of correlation.
MATH M15H-use the chi-square distribution to test independence and to test goodness of fit.
MATH M15H-conduct a one-way Analysis of Variance (ANOVA) hypothesis test.

MATH M15H-select an appropriate hypothesis test and interpret the result using p-value; use appropriate statistical technique to analyze and interpret applications based on data related to business, social sciences, psychology, life sciences, health sciences or education, and interpret results using technology-based statistical analysis.
MATH M15H-Honors: find probabilities using Poisson distributions.
MATH M15H-Honors: estimate binomial probabilities.
MATH M15H-Honors: find and use standard deviations.
MATH M25C-apply the basic rules of vector algebra to carry out vector operations in the plane and in space.
MATH M25C-evaluate dot, cross and triple scalar products and projections.
MATH M25C-use the dot product, cross product, and triple scalar product to solve applied problems.
MATH M25C-write the parametric equations and symmetric equations of a line in space and write the standard equation of a plane in space.
MATH M25C-identify the six quadric surfaces.
MATH M25C-convert from rectangular to the cylindrical and spherical coordinates in space.
MATH M25C-evaluate derivatives and integrals of vector-valued functions.
MATH M25C-compute velocity and acceleration vectors for vector-valued functions.
MATH M25C-find the tangential and normal components of an acceleration vector and compute arc length and curvature of a space curve.
MATH M25C-evaluate limits and determine continuity for functions of two variables at a point.
MATH M25C-find the first-order and higher-order partial derivatives of functions of several variables, determine differentiability, and apply the chain rule to find partial derivatives.
MATH M25C-compute the directional derivative and the gradient vector for a function of two or three variables.
MATH M25C-write the equation of a tangent plane at a point on a surface.
MATH M25C-find and classify all critical points for a function of two variables and use Lagrange multipliers to find maxima and minima of functions of two variables subject to side conditions.
MATH M25C-use double integrals to compute areas and volumes and surface areas.
MATH M25C-evaluate double integrals using polar coordinates.
MATH M25C-find the center of mass of a variable density planar lamina.
MATH M25C-evaluate triple integrals using rectangular, cylindrical, or spherical coordinates.
MATH M25C-compute the potential function, curl, and divergence of a vector field.
MATH M25C-evaluate the line integral of a vector field on a curve and surface integrals.
MATH M25C-apply Green's Theorem to compute line integrals in the plane.
MATH M25C-use the Divergence Theorem to compute the flux of a vector field through a surface.
MATH M25C-use Stokes's Theorem to compute line integrals within a vector field around a closed curve.
MATH M31-solve systems of linear equations using various methods including Gaussian and Gauss-Jordan elimination and inverse matrices.
MATH M31-perform matrix algebra, invertibility, and the transpose and understand vector algebra in R n .
MATH M31-determine relationship between coefficient matrix invertibility and solutions to a system of linear equations and the inverse matrices.
MATH M31-define special matrices: diagonal, triangular, and symmetric.
MATH M31-understand determinants and their properties.
MATH M31-understand real vector spaces and subspaces and apply their properties.
MATH M31- understand linear independence and dependence.
MATH M31-find basis and dimension of a vector space,and understand change of basis.
MATH M31-find a basis for the row space, column space and null space of a matrix and find the rank and nullity of a matrix.
MATH M31-compute linear transformations, kernel and range, and inverse linear transformations, and find matrices of general linear transformations.
MATH M31-find the dimension of spaces such as those associated with matrices and linear transformations.
MATH M31-find eigenvalues and eigenvectors and use them in applications.
MATH M31-diagonalize and orthogonally diagonalize symmetric matrices.
MATH M31-evaluate the dot product, norm, angle between vectors, and orthogonality of two vectors in R n .
MATH M31-compute inner products on a real vector space and compute angle and orthogonality in inner product spaces.
MATH M31-create orthogonal and orthonormal bases: Gram-Schmidt process and use bases and orthonormal bases to solve application problems.
MATH M31-prove basic results in linear algebra using appropriate proof-writing techniques such as linear independence of vectors; properties of subspaces; linearity, injectivity and surjectivity of functions; and properties of eigenvectors and eigenvalues.

Requisite Justification
Requisite Type
Prerequisite

Requisite
MATH M15 OR MATH M15H

Requisite Description
Course not in a sequence

Level of Scrutiny/Justification
Required by 4 year institution

Requisite Type
Prerequisite

Requisite
MATH M25C

Requisite Description
Course not in a sequence

Level of Scrutiny/Justification
Required by 4 year institution

Requisite Type
Prerequisite

Requisite
MATH M31

Requisite Description
Course not in a sequence

Level of Scrutiny/Justification
Required by 4 year institution

Requisite Type
Prerequisite

Requisite
CS M10DS

Requisite Description
Course not in a sequence

Level of Scrutiny/Justification
Required by 4 year institution

Requisite Type
Recommended Preparation

Requisite
MATH M21

Requisite Description
Course not in a sequence

Level of Scrutiny/Justification
Required by 4 year institution

Requisite Type
Recommended Preparation

Requisite
MATH M35

Requisite Description
Course not in a sequence

Level of Scrutiny/Justification
Required by 4 year institution

Student Learning Outcomes (CSLOs)

Upon satisfactory completion of the course, students will be able to:

1	explain how a matrix can be used to store data and be manipulated with a programming language (e.g., Python).
2	compute a metric.
3	explain the similarities and differences between the gradient descent, stochastic gradient descent and ADAM algorithms.

Course Objectives

Upon satisfactory completion of the course, students will be able to:

1	solve a system of linear equations with matrices.
2	apply special matrices (such as positive definite or symmetric) to data sets.
3	perform matrix arithmetic, matrix operations (e.g., determinant, transpose, dot product.)
4	determine the basis, span, rank, and other properties of vector spaces.
5	find the inverse of a matrix.
6	find the four fundamental subspaces (column, row, null, and left nullspace.)
7	perform linear/matrix transformations.
8	find orthonormal bases, orthogonal complement, and orthogonal projections of vector spaces.
9	compute the eigenvalues and eigenvectors of a matrix.
10	perform matrix decompositions such as eigendecomposition, singular value decomposition (SVD), and non-negative matrix factorization (NMF).
11	find the following metrics: the L1 norm (Manhattan Distance), L2 norm (Euclidian distance), and Minkowski distance.
12	apply kernel functions such as polynomial, Gaussian, Gaussian RBF, Laplace RBF, hyperbolic tangent, and sigmoid.
13	perform Principal Component Analysis (PCA.)
14	perform constrained optimization with Lagrange multipliers.
15	use the gradient descent, stochastic gradient descent and ADAM algorithms.
16	apply the backpropagation algorithm and understand how its related to the chain rule.
17	computing and interpreting the ROC curves and confusion matrix from data.
18	recognize the ethical issues with data.
19	use a programming language such as Python to perform calculations.

Course Content

Lecture/Course Content

15% Applied Linear Algebra

1. System of linear equations and special matrices
2. Matrix arithmetic and operations
3. Properties of vector spaces and their subspaces
4. Linear transformations

20% Analytic Geometry

1. Norms
2. Orthonormal/orthogonal properties of vector spaces

20% Matrix Decompositions

1. Eigenvalues and Eigenvectors
2. Matrix Decompositions

20% Vector Calculus & Optimization

1. Constrained optimization
2. Gradient descent
3. Backpropagation

20% Dimensionality & Classification

1. Kernels
2. PCA
3. Confusion Matrix and ROC curves

5% Ethics

1. Ethics and data

Laboratory or Activity Content

None.

Methods of Evaluation

Which of these methods will students use to demonstrate proficiency in the subject matter of this course? (Check all that apply):

Problem solving exercises
Written expression

Methods of Evaluation may include, but are not limited to, the following typical classroom assessment techniques/required assignments (check as many as are deemed appropriate):

Computational homework
Group projects
Individual projects
Mathematical proofs
Objective exams
Oral presentations
Problem-solving exams
Quizzes

Instructional Methodology

Specify the methods of instruction that may be employed in this course

Computer-aided presentations
Collaborative group work
Class activities
Class discussions
Distance Education
Guest speakers
Instructor-guided interpretation and analysis
Instructor-guided use of technology
Lecture
Small group activities

Describe specific examples of the methods the instructor will use:

- Using whiteboard, iPad, Zoom, or other technology to lecture on the difference and similarities of different types of distance metrics.
- Classroom discussion, with student response, such as discussing when to use a particular type of algorithm.
- In-class activities where students work in small groups to determine the perform PCA.
- Instructor demonstrating how to use a programming language to perform operation on matrices.

Representative Course Assignments

Writing Assignments

- Homework problems selected from the textbook where answers require a written explanation of the solution such computing the Euclidian distance.
- A machine learning scenario on an exam where the student must choose the correct machine learning algorithm for that scenario.
- Graded in-class and/or homework assignments requiring the use of a programming language where the student needs to write the comments for the code and interpret the results using written English.

Critical Thinking Assignments
- Elucidate the moral issues that arise with the use of data orally.
- Explicate a confusion matrix.
- Interpret the computer generated output from a computer language that is being applied to machine learning.

Reading Assignments
- Read and understand the help documents for the Python programming language used in the class.
- Read scholarly and/or news articles on using data for the social good.

Skills Demonstrations
- Use a programming language such as Python, to find the ROC curve for an algorithm and interpret it.
- Find the four fundamental subspaces of vector space.
- Additional problem sets by the instructor such as those using optimization with Lagrange multipliers.

Outside Assignments

Representative Outside Assignments
- Group or individual linear algebra project, such as performing machine learning on a large data set with missing data and interpreting the mathematics behind the algorithm.
- Graded code problems, such as using code to perform machine learning algorithm.
- Assign a reading on the use and misuse of data by Big Tech.

Articulation

Equivalent Courses at 4 year institutions

University	Course ID	Course Title	Units
UC Berkeley	STAT 89A	Linear Algebra for Data Science	4
UC Los Angeles	MATH 42	Introduction to Data-Driven Mathematical Modeling: Life, The Universe, and Everything	4
UC San Diego	DSC 40A	Theoretical Foundations of Data Science I	4

District General Education

A. Natural Sciences

B. Social and Behavioral Sciences

C. Humanities

D. Language and Rationality

E. Health and Physical Education/Kinesiology

F. Ethnic Studies/Gender Studies

Course is CSU transferable
Yes

CSU Baccalaureate List effective term:
S2022

CSU GE-Breadth

Area A: English Language Communication and Critical Thinking

Area B: Scientific Inquiry and Quantitative Reasoning

Area C: Arts and Humanities

Area D: Social Sciences

Area E: Lifelong Learning and Self-Development

Area F: Ethnic Studies

CSU Graduation Requirement in U.S. History, Constitution and American Ideals:

UC TCA

UC TCA
Proposed

Date Proposed:
6/15/2021

IGETC

Area 1: English Communication

Area 2A: Mathematical Concepts & Quantitative Reasoning

Area 3: Arts and Humanities

Area 4: Social and Behavioral Sciences

Area 5: Physical and Biological Sciences

Area 6: Languages Other than English (LOTE)

Textbooks and Lab Manuals

Resource Type
Textbook

Classic Textbook
No

Description
Deisenroth, Marc Peter, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for Machine Learning*. Cambridge UP, 2020.

Resource Type
Software

Description
Anaconda Navigator, includes both Jupyter Notebook for Python and R Studio for R. https://docs.anaconda.com/anaconda/navigator/

Library Resources

Assignments requiring library resources

Research using the library's print and online resources to locate data for projects. If the student does not have their own computer, then possible use of the Library's for programming.

Sufficient Library Resources exist

Yes

Example of Assignments Requiring Library Resources

Using Library's print, online, and computer resources to locate scholarly and/or news articles on using data for the social good.

Distance Education Addendum

Definitions

Distance Education Modalities

Hybrid (51%–99% online)
Hybrid (1%–50% online)
100% online

Faculty Certifications

Faculty assigned to teach Hybrid or Fully Online sections of this course will receive training in how to satisfy the Federal and state regulations governing regular effective/substantive contact for distance education. The training will include common elements in the district-supported learning management system (LMS), online teaching methods, regular effective/substantive contact, and best practices.

Yes

Faculty assigned to teach Hybrid or Fully Online sections of this course will meet with the EAC Alternate Media Specialist to ensure that the course content meets the required Federal and state accessibility standards for access by students with disabilities. Common areas for discussion include accessibility of PDF files, images, captioning of videos, Power Point presentations, math and scientific notation, and ensuring the use of style mark-up in Word documents.

Yes

Regular Effective/Substantive Contact

Hybrid (1%–50% online) Modality:

Method of Instruction	Document typical activities or assignments for each method of instruction
Asynchronous Dialog (e.g., discussion board)	Use of student discussion boards to discuss concepts from the material, solutions to homework problems, general discussion of techniques in solving problems, study skills, or arranging study groups.
E-mail	Responding to student queries about material, grade information, course policies and procedures, scheduling and due dates, submitting homework assignments, or making general announcements to the class.
Face to Face (by student request; cannot be required)	Students requesting to speak to instructor in person for personal help on material, grade information, or discussion of policies and procedures.
Other DE (e.g., recorded lectures)	Posting of recorded lectures either by the instructor, recorded lessons available through campus resources, or use of public online resources available on the internet.
Synchronous Dialog (e.g., online chat)	Active live discussion with the instructor on material concepts, techniques for problem solving, feedback on solutions to problems, general chat on study skills, or answers to homework problems, quizzes or tests.

Hybrid (51%–99% online) Modality:

Method of Instruction	Document typical activities or assignments for each method of instruction
Asynchronous Dialog (e.g., discussion board)	Use of student discussion boards to discuss concepts from the material, solutions to homework problems, general discussion of techniques in solving problems, study skills, or arranging study groups.

E-mail	Responding to student queries about material, grade information, course policies and procedures, scheduling and due dates, submitting homework assignments, or making general announcements to the class.
Face to Face (by student request; cannot be required)	Students requesting to speak to instructor in person for personal help on material, grade information, or discussion of policies and procedures.
Other DE (e.g., recorded lectures)	Posting of recorded lectures either by the instructor, recorded lessons available through campus resources, or use of public online resources available on the internet.
Synchronous Dialog (e.g., online chat)	Active live discussion with the instructor on material concepts, techniques for problem solving, feedback on solutions to problems, general chat on study skills, or answers to homework problems, quizzes or tests.

100% online Modality:

Method of Instruction	Document typical activities or assignments for each method of instruction
Asynchronous Dialog (e.g., discussion board)	Use of student discussion boards to discuss concepts from the material, solutions to homework problems, general discussion of techniques in solving problems, study skills, or arranging study groups.
E-mail	Responding to student queries about material, grade information, course policies and procedures, scheduling and due dates, submitting homework assignments, or making general announcements to the class.
Face to Face (by student request; cannot be required)	Students requesting to speak to instructor in person for personal help on material, grade information, or discussion of policies and procedures.
Other DE (e.g., recorded lectures)	Posting of recorded lectures either by the instructor, recorded lessons available through campus resources, or use of public online resources available on the internet.
Synchronous Dialog (e.g., online chat)	Active live discussion with the instructor on material concepts, techniques for problem solving, feedback on solutions to problems, general chat on study skills, or answers to homework problems, quizzes or tests.

Examinations

Hybrid (1%–50% online) Modality
Online
On campus

Hybrid (51%–99% online) Modality
Online
On campus

Primary Minimum Qualification
MATHEMATICS

Review and Approval Dates

Department Chair
03/16/2021

Dean
03/16/2021

Technical Review
03/25/2021

Curriculum Committee
04/06/2021

DTRW-I
04/08/2021

Curriculum Committee
MM/DD/YYYY

Board
05/11/2021

CCCCO
MM/DD/YYYY

DOE/accreditation approval date
MM/DD/YYYY