DSC 305 - Machine Learning - Spring 2025 Syllabus

Marcus Birkenkrahe

January 6, 2025

1 General Course Information

- Meeting Times: Tuesday/Thursday, 14:30-15:45 hrs
- Meeting place: Lyon Building Computer Lab 104
- Professor: Dr. Marcus Birkenkrahe
- Office: Derby Science Building 210
- Phone: (870) 307-7254
- Office Hours: By appointment via tinyurl.com/sp25-booking
- Textbook: Machine Learning With R (3e) by Brett Lantz, Packt Publishing, 2019 (Online: packtpub.com, ebook); & DataCamp

2 Standard and course policies

- **Standard Lyon College Policies** are incorporated into this syllabus and can be found at: lyon.edu/standard-course-policies.
- The **Assignments and Honor Code** and the **Attendance Policy** are incorporated into this syllabus also and can be found at: tinyurl.com/LyonPolicy.
- In addition to these rules, please read and observe my guide to Using AI to code (written in Fall 2024): tinyurl.com/Using-AI-to-code.

3 Objectives

This course is concerned with algorithms that transform information into actionable intelligence using present-day computers and big data. We use R, a cross-platform, zero-cost statistical programming environment that combines a wide range of functions, interfaces to common machine learning packages, and best-in-class visualization.

4 Student learning outcomes

Students who complete DSC 305, "Machine Learning" (ML), can:

- **Understand core machine learning concepts** including data exploration, classification, regression, and clustering techniques.
- **Apply machine learning algorithms** such as k-NN, Naive Bayes, decision trees, regression, neural networks, and support vector machines using R.
- **Evaluate and improve model performance** using appropriate metrics and strategies to optimize predictive accuracy.
- **Analyze real-world datasets** by implementing machine learning solutions to classification, regression, and clustering problems.
- **Interpret patterns and insights** from data through methods like market basket analysis and k-means clustering.
- **Manage and preprocess data** to ensure accuracy and relevance for machine learning models.
- Know how to effectively **present assignment and project results**

5 Course requirements

- **Basic proficiency with R** is useful (as taught in DSC 105 or obtained independently on DataCamp "Introduction to R", GitHub's "fasteR", or Part I of Book of R by Davies, NoStarch 2016).
- Imagination, creativity and a visual mind, enjoying finding patterns and spotting correlations.

- Basic understanding of algorithms and data structures (in any programming language).
- Basic understanding of data science infrastructure especially literate programming methods.

6 Grading system

You should be able to see your current grade at any time using the Canvas gradebook for the course.

WHEN	DESCRIPTION	IMPACT
Weekly	Assignments	25%
Weekly	Multiple choice tests	25%
Monthly	Project Sprint Reviews	25%
TBD	Final exam (optional)	25%

Notes:

- **To pass:** 60% of all available points.
- **Tests:** weekly online quizzes based on classroom lectures and practice.
- **Final exam:** random selection of the known test questions. **Note:** You only have to write the final exam if you want to improve your grade at the end of the course. If the final exam result is below your final grade average up to this point, it will be ignored.

7 Rubric

Component	Weight	Excellent	Good	Satisfactory	Needs Improvement	Unsatisfactory
Participation and Attendance	0%	Consistently attends and actively participates in all classes.	Attends most classes and participates in discussions.	Attends classes but participation is minimal.	Frequently absent and rarely participates.	Rarely attends classes and does not participate.
Programming assignments	50%	Completes all assignments on time with high accuracy (90-100%).	Completes most assignments on time with good accuracy (80-89%).	Completes assignments but with some inaccuracies or delays (70-79%).	Frequently late or incomplete assignments with several inaccuracies (60-69%).	Rarely completes assignments and shows minimal understanding (0-59%).
Tests	25%	Demonstrates thorough understanding and application of concepts (90-100%).	Shows good understanding with minor errors (80-89%).	Displays basic understanding with some errors (70-79%).	Limited understanding with several errors (60-69%).	Minimal understanding and many errors (0-59%).
Final Exam (Optional)	25%	Demonstrates comprehensive understanding and application of course concepts (90-100%).	Shows strong understanding with minor errors (80-89%).	Displays adequate understanding with some errors (70-79%).	Limited understanding with several errors (60-69%).	Minimal understanding and many errors (0-59%).

::

8 Grading table

This table is used to convert completion rates into letter grades. for the midterm results, letter grades still carry signs, while for the term results, only straight letters are given (by rounding up).

Percentage	LETTER GRADE
100% to 89.5%	A (very good)
$<$ 89.5% to 79.5%	B (good)
$<$ 79.5% to 69.5%	C (satisfactory)
$<$ 69.5% to 59.5%	D (passed)
$<$ 59.5% to 0%	F (FAILED)

9 Schedule and session content

For **important dates**, see the 2024-2025 Academic Calendar at: catalog.lyon.edu/202425-academic-calendar

Week	Topic
1	What is Machine Learning?
2	Machine Learning Models
3	Lazy Learning – Classification Using Nearest Neighbors
4	Probabilistic Learning – Classification Using Naive Bayes
5	Divide and Conquer – Classification Using Decision Trees
6	Divide and Conquer – Classification Using Rules
7	Forecasting Numeric Data – Regression Methods
8	Regression Trees and Model Trees
9	Finding Patterns – Market Basket Analysis
10	Finding Groups of Data – Clustering with k-means
11	Unsupervised Learning
12	Hierarchical Clustering
13	Dimensionality reduction with PCA
14	Case studies and course review
15	Project Presentations
16	Final exam week