Data Mining

Course Overview and Logistics

https://data-mining.github.io/winter-2026/

CS 453/553 – Winter 2026

Yu Wang, Ph.D.
Assistant Professor
Computer Science
University of Oregon

Course Logistics – Quzz

Times:

- **Classes:** Monday/Wednesday 12:00-1:20 pm PST, Gerlinger 302
- **Office hours:** Wednesday 1:20-2:00 pm PST, other time by appointment
- **Zoom:** https://uoregon.zoom.us/j/4052006678

Components:

Course Assessment and Grading Scale

Category	CS-453 (%)	CS-553 (%)
Quzz 1	20%	15%
Quzz 2	20%	15%
Project	40%	45%
Participation	5%	5%
Paper Presentation	15%	20%
Overleaf Bonus	5%	5%

- As long as you are **active thinking** and **understand the content**, you will be good

Course Logistics – Project

Times:

- **Classes:** Monday/Wednesday 12:00-1:20 pm PST, Gerlinger 302
- **Office hours:** Wednesday 1:20-2:00 pm PST, other time by appointment
- **Zoom:** https://uoregon.zoom.us/j/4052006678

Components:

Course Assessment and Grading Scale

Category	CS-453 (%)	CS-553 (%)
Quzz 1	20%	15%
Quzz 2	20%	15%
Project	40%	45%
Participation	5%	5%
Paper Presentation	15%	20%
Overleaf Bonus	5%	5%

https://ml-graph.github.io/winter-2026/project/

Course Logistics – Project

Project

The project may be completed either individually or as a team; both approaches are acceptable. For team-based projects, only one team member should submit the final report and clearly specify all contributing teammates. Bonus Points will apply if you consider doing projects in the following fields with (*) or any domain beyond the following:

1. Background and Problem Formulation - 10%

- **Background - 5%:**
  - What is the general background of the problem you are working on?
    - I want to develop a better paper categorization system
- **Problem Formulation - 5%:**
  - Under the general topic, what specific problem is your project addressing?
    - I want to develop a machine learning model/algorithm to take input of the paper, output the paper topic (machine learning, computer system, human-computer collaboration, etc.)

2. Data Mining Stage - 35%

- **Data Collection and Store - 15%:**
  - What data are you looking to kick off your project? How do you collect them? What data structure do you use to represent them?
    - I collect Cora/Citeseer/Pubmed Data from somewhere (e.g., a paper, a GitHub repository, Hugging Face, etc.), and I use an adjacency list to store their connection and a matrix to store their node feature
- **Data Mining - 20%:**
  - What kind of data mining problem do you need to do and why?
    - I need to analyze the network homophily/heterophily since leveraging this property might help me develop a better machine learning model for paper classification.
  - How do you do it?
    - I calculate for every edge, the two ending points, whether they are in the same class or not, and quantify the average ratio as a homophily ratio
  - What kind of pattern do you find? How do you present your findings/analysis?
    - I find that in many paper citation networks, the homophily is pretty high. Using Number/Table/Figure, etc.

3. Machine Learning Stage - 35%

- **Machine Learning Model Design:**
  - Based on your targeted problem, what kind of machine learning model do you want to build and why?
    - I want to build a graph neural network to fully exploit the discovered homophily principle.

https://ml-graph.github.io/winter-2026/project/

Course Logistics – Paper Presentation

Times:

- **Classes:** Monday/Wednesday 12:00-1:20 pm PST, Gerlinger 302
- **Office hours:** Wednesday 1:20-2:00 pm PST, other time by appointment
- **Zoom:** https://uoregon.zoom.us/j/4052006678

Components:

Course Assessment and Grading Scale

Category	CS-453 (%)	CS-553 (%)
Quzz 1	20%	15%
Quzz 2	20%	15%
Project	40%	45%
Participation	5%	5%
Paper Presentation	15%	20%
Overleaf Bonus	5%	5%

https://ml-graph.github.io/winter-2026/presentation/

Course Logistics – Paper Presentation

Presentation

Paper Presentation Details

You can either collaborate with a team or present individually. The choice of topic is entirely up to you.

- Introduction and Background – What is the general impact and background of the topic?
- Motivation and Problem – What is the core research problem, and why do we study it?
- Related Work and Challenges – How did previous works address this problem, and what are some of the challenges?
- Proposed Solutions/Methods and Rationale – What are the proposed methods/techniques, and why are they proposed? What specific reasons would solving this problem require these proposed(1) methods/techniques?
- Experimental Setting, Results, and Analysis – What experiments are designed to verify the proposed method? How are results being discussed and analyzed? Are there any interesting findings?
- Conclusion and Future Work

Do not use sentences in the slides, but use bullet points and important points that you can logically chain together for your speech I will be very careful taking note of this. Please pardon me for this!

Natural Disaster Modeling

Neural-Biology Analysis

Social Network

Agentic AI

Reasoning/Planning

Knowledge Retrieval

https://ml-graph.github.io/winter-2026/presentation/

Course Logistics – Paper Presentation

Presentation

Paper Presentation Details

You can either collaborate with a team or present individually. The choice of topic is entirely up to you.

- Introduction and Background – What is the general impact and background of the topic?
- Motivation and Problem – What is the core research problem, and why do we study it?
- Related Work and Challenges – How did previous works address this problem, and what are some of the challenges?
- Proposed Solutions/Methods and Rationale – What are the proposed methods/techniques, and why are they proposed? What specific reasons would solving this problem require these proposed(1) methods/techniques?
- Experimental Setting, Results, and Analysis – What experiments are designed to verify the proposed method? How are results being discussed and analyzed? Are there any interesting findings?
- Conclusion and Future Work

Do not use sentences in the slides, but use bullet points and important points that you can logically chain together for your speech I will be very careful taking note of this. Please pardon me for this!

Natural Disaster Modeling

Neural-Biology Analysis

Social Network

Agentic AI

Reasoning/Planning

https://ml-graph.github.io/winter-2026/presentation/

Course Logistics – Paper Presentation – Bad Example

The provided image outlines the logistical and academic requirements for a course at the University of Oregon, likely **CS-453/553**. Classes are held on **Mondays and Wednesdays from 12:00 pm to 1:20 pm PST** in Gerlinger 302, with office hours scheduled for Wednesdays from 1:20 pm to 2:00 pm or by appointment. A specific Zoom link is also provided for virtual access.

The grading structure, labeled "Course Assessment and Grading Scale," distinguishes between undergraduate (**CS-453**) and graduate (**CS-553**) requirements. For undergraduate students, the grade is heavily weighted toward two quizzes at **20% each** (40% total) and a project worth **40%**, followed by a paper presentation at **15%** and participation at **5%**. Graduate students have a slightly different distribution, with quizzes weighted less at **15% each** (30% total), while the project and paper presentation are weighted higher at **45%** and **20%** respectively.

Both groups have the opportunity for a **5% Overleaf Bonus**. Beside the grading chart, a motivational note emphasizes that students will succeed as long as they maintain **active thinking** and **understand the content**.

Course Logistics – Paper Presentation – Good Example

Times:

- **Classes:** Monday/Wednesday 12:00-1:20 pm PST, Gerlinger 302
- **Office hours:** Wednesday 1:20-2:00 pm PST, other time by appointment
- **Zoom:** https://uoregon.zoom.us/j/4052006678

Components:

Course Assessment and Grading Scale

Category	CS-453 (%)	CS-553 (%)
Quzz 1	20%	15%
Quzz 2	20%	15%
Project	40%	45%
Participation	5%	5%
Paper Presentation	15%	20%
Overleaf Bonus	5%	5%

- As long as you are **active thinking** and **understand the content**, you will be good

Course Logistics – Timeline

Basics

EVENT	DATE	DESCRIPTION	COURSE MATERIAL
Lecture	01/05/2026 Monday	Overview Syllabus	Course Materials: ◦ Slides
Assignment	01/05/2026 Monday	**Project released!**	[Project]
Lecture	01/07/2026 Wednesday	Logistics Basics	Course Materials: ◦ Slides
Lecture	01/12/2026 Monday	Classification KNN/Naive Bayes	Course Materials: ◦ Slides
Lecture	01/14/2026 Wednesday	Classification Decision Tree	Course Materials: ◦ Slides
Martin Luther King, Jr holiday	01/19/2026 04:30 Monday	Enjoy :)	
Lecture	01/21/2026 Wednesday	Clustering K-means, Hierarchical Clustering	Course Materials: ◦ Slides
Lecture	01/26/2026 Monday	Dimension Reduction PCA	Course Materials: ◦ Slides
Lecture	01/28/2026 Wednesday	Linear Regression Gradient Descent	Course Materials: ◦ Slides
Lecture	02/02/2026 Monday	Logistic Classification	Course Materials: ◦ Slides
Lecture	02/04/2026 Wednesday	Neural Network	Course Materials: ◦ Slides
Exam	**02/09/2026 16:00 Monday**	**Quzz 1**	Topics: ◦ Lecture 1 - Lecture 8 ◦ Closed Book

Advanced

Lecture	02/11/2026 Wednesday	Graph Mining	Course Materials: ◦ Slides
	02/11/2026 16:00 Wednesday	Presentation 1	Group ◦ Group 1: 7-7:15 pm ◦ Group 2: 7:15-7:30 pm ◦ Zoom
Lecture	02/16/2026 Monday	Graph Mining	Course Materials: ◦ Slides
	02/16/2026 16:00 Monday	Presentation 2	Group ◦ Group 3: 7-7:15 pm ◦ Group 4: 7:15-7:30 pm ◦ Zoom
Lecture	02/18/2026 Wednesday	Temporal Mining	Course Materials: ◦ Slides
	02/18/2026 16:00 Wednesday	Presentation 3	Group ◦ Group 5: 7-7:15 pm ◦ Group 6: 7:15-7:30 pm ◦ Zoom
Lecture	02/23/2026 Monday	Spatial Cloud Point Mining	Course Materials: ◦ Video Record ◦ Slides ◦ Video
	02/23/2026 16:00 Monday	Presentation 4	Group ◦ Group 7: 7-7:15 pm ◦ Group 8: 7:15-7:30 pm ◦ Zoom
Lecture	02/25/2026 Wednesday	Image Mining	Course Materials: ◦ Video Record ◦ Slides ◦ Video
	02/25/2026 16:00 Wednesday	Presentation 5	Group ◦ Group 9: 7-7:15 pm ◦ Group 10: 7:15-7:30 pm ◦ Zoom
Lecture	03/02/2026 Monday	Language Mining	Course Materials: ◦ Slides
	03/02/2026 16:00 Monday	Presentation 6	Group ◦ Group 11: 7-7:15 pm ◦ Group 12: 7:15-7:30 pm ◦ Zoom
Lecture	03/04/2026 Wednesday	Language Mining	Course Materials: ◦ Slides
	03/04/2026 16:00 Wednesday	Presentation 6	Group ◦ Group 13: 7-7:15 pm ◦ Group 14: 7:15-7:30 pm ◦ Zoom
Lecture	03/09/2026 Monday	Review Future	Course Materials: ◦ Slides
	03/09/2026 16:00 Monday	Presentation 6	Group ◦ Group 15: 7-7:15 pm ◦ Group 16: 7:15-7:30 pm ◦ Zoom
Exam	**03/11/2026 16:00 Wednesday**	**Quzz 2**	Topics: ◦ Lecture 9 - Lecture 16 ◦ Closed Book
Due	**03/20/2026 23:59 Friday**	**Project Report Due**	

Phase 1 + Quzz 1

Phase 2 + Quzz 2 + Project Report

Course Logistics – Timeline

Basics

EVENT	DATE	DESCRIPTION	COURSE MATERIAL
Lecture	01/05/2026 Monday	Overview Syllabus	Course Materials: Slides
Assignment	01/05/2026 Monday	**Project released!**	[Project]
Lecture	01/07/2026 Wednesday	Logistics Basics	Course Materials: Slides
Lecture	01/12/2026 Monday	Classification KNN/Naive Bayes	Course Materials: Slides
Lecture	01/14/2026 Wednesday	Classification Decision Tree	Course Materials: Slides
Martin Luther King, Jr holiday	01/19/2026 04:30 Monday	Enjoy :)	
Lecture	01/21/2026 Wednesday	Clustering K-means, Hierarchical Clustering	Course Materials: Slides
Lecture	01/26/2026 Monday	Dimension Reduction PCA	Course Materials: Slides
Lecture	01/28/2026 Wednesday	Linear Regression Gradient Descent	Course Materials: Slides
Lecture	02/02/2026 Monday	Logistic Classification	Course Materials: Slides
Lecture	02/04/2026 Wednesday	Neural Network	Course Materials: Slides
Exam	**02/09/2026 16:00 Monday**	**Quzz 1**	Topics: Lecture 1 - Lecture 8; Closed Book

Advanced

Lecture	02/11/2026 Wednesday	Graph Mining	Course Materials: Slides
	02/11/2026 16:00 Wednesday	Presentation 1	Group: Group 1: 7-7:15 pm; Group 2: 7:15-7:30 pm; Zoom
Lecture	02/16/2026 Monday	Graph Mining	Course Materials: Slides
	02/16/2026 16:00 Monday	Presentation 2	Group: Group 3: 7-7:15 pm; Group 4: 7:15-7:30 pm; Zoom
Lecture	02/18/2026 Wednesday	Temporal Mining	Course Materials: Slides
	02/18/2026 16:00 Wednesday	Presentation 3	Group: Group 5: 7-7:15 pm; Group 6: 7:15-7:30 pm; Zoom
Lecture	02/23/2026 Monday	Spatial Cloud Point Mining	Course Materials: Video Record; Slides; Video
	02/23/2026 16:00 Monday	Presentation 4	Group: Group 7: 7-7:15 pm; Group 8: 7:15-7:30 pm; Zoom
Lecture	02/25/2026 Wednesday	Image Mining	Course Materials: Video Record; Slides; Video

	02/25/2026 16:00 Wednesday	Presentation 5	Group: Group 9: 7-7:15 pm; Group 10: 7:15-7:30 pm; Zoom
Lecture	03/02/2026 Monday	Language Mining	Course Materials: Slides
	03/02/2026 16:00 Monday	Presentation 6	Group: Group 11: 7-7:15 pm; Group 12: 7:15-7:30 pm; Zoom
Lecture	03/04/2026 Wednesday	Language Mining	Course Materials: Slides
	03/04/2026 16:00 Wednesday	Presentation 6	Group: Group 13: 7-7:15 pm; Group 14: 7:15-7:30 pm; Zoom
Lecture	03/09/2026 Monday	Review Future	Course Materials: Slides
	03/09/2026 16:00 Monday	Presentation 6	Group: Group 15: 7-7:15 pm; Group 16: 7:15-7:30 pm; Zoom
Exam	**03/11/2026 16:00 Wednesday**	**Quzz 2**	Topics: Lecture 9 - Lecture 16; Closed Book
Due	**03/20/2026 23:59 Friday**	**Project Report Due**	

Paper Presentation

Phase 1 + Quzz 1

Phase 2 + Quzz 2 + Project Report

Course Logistics – Timeline

Basics

EVENT	DATE	DESCRIPTION	COURSE MATERIAL
Lecture	01/05/2026 Monday	Overview Syllabus	Course Materials: Slides
Assignment	01/05/2026 Monday	**Project released!**	[Project]
Lecture	01/07/2026 Wednesday	Logistics Basics	Course Materials: Slides
Lecture	01/12/2026 Monday	Classification KNN/Naive Bayes	Course Materials: Slides
Lecture	01/14/2026 Wednesday	Classification Decision Tree	Course Materials: Slides
Martin Luther King, Jr holiday	01/19/2026 04:30 Monday	Enjoy :)	
Lecture	01/21/2026 Wednesday	Clustering K-means, Hierarchical Clustering	Course Materials: Slides
Lecture	01/26/2026 Monday	Dimension Reduction PCA	Course Materials: Slides
Lecture	01/28/2026 Wednesday	Linear Regression Gradient Descent	Course Materials: Slides
Lecture	02/02/2026 Monday	Logistic Classification	Course Materials: Slides
Lecture	02/04/2026 Wednesday	Neural Network	Course Materials: Slides
Exam	**02/09/2026 16:00 Monday**	**Quzz 1**	Topics: Lecture 1 - Lecture 8; Closed Book

Advanced

Out of Town, Video Record

Lecture	02/11/2026 Wednesday	Graph Mining	Course Materials: Slides
	02/11/2026 16:00 Wednesday	Presentation 1	Group: Group 1: 7-7:15 pm; Group 2: 7:15-7:30 pm; Zoom
Lecture	02/16/2026 Monday	Graph Mining	Course Materials: Slides
	02/16/2026 16:00 Monday	Presentation 2	Group: Group 3: 7-7:15 pm; Group 4: 7:15-7:30 pm; Zoom
Lecture	02/18/2026 Wednesday	Temporal Mining	Course Materials: Slides
	02/1[illegible] 16:00 Wednesday	Presentation 3	Group: Group 5: 7-7:15 pm; Group 6: 7:15-7:30 pm; Zoom
Lecture	02/23/2026 Monday	Spatial Cloud Point Mining	Course Materials: Video Record; Slides; Video
	02/23/2026 16:00 Monday	Presentation 4	Group: Group 7: 7-7:15 pm; Group 8: 7:15-7:30 pm; Zoom
Lecture	02/25/2026 Wednesday	Image Mining	Course Materials: Video Record; Slides; Video

	02/25/2026 16:00 Wednesday	Presentation 5	Group: Group 9: 7-7:15 pm; Group 10: 7:15-7:30 pm; Zoom
Lecture	03/02/2026 Monday	Language Mining	Course Materials: Slides
	03/02/2026 16:00 Monday	Presentation 6	Group: Group 11: 7-7:15 pm; Group 12: 7:15-7:30 pm; Zoom
Lecture	03/04/2026 Wednesday	Language Mining	Course Materials: Slides
	03/04/2026 16:00 Wednesday	Presentation 6	Group: Group 13: 7-7:15 pm; Group 14: 7:15-7:30 pm; Zoom
Lecture	03/09/2026 Monday	Review Future	Course Materials: Slides
	03/09/2026 16:00 Monday	Presentation 6	Group: Group 15: 7-7:15 pm; Group 16: 7:15-7:30 pm; Zoom
Exam	**03/11/2026 16:00 Wednesday**	**Quzz 2**	Topics: Lecture 9 - Lecture 16; Closed Book
Due	**03/20/2026 23:59 Friday**	**Project Report Due**	

Phase 1 + Quzz 1

Phase 2 + Quzz 2 + Project Report

Question Time!

- **Linear Algebra**

- **Statistics/Probability**

Vector

$$\mathbf{v} = \begin{bmatrix} 1 & 2 & 5 \end{bmatrix}$$

$$\boxed{\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}}$$

Please note that we will use this one by default

Matrix

$$\mathbf{A} = \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix}}_{\text{3 columns}} \left.\vphantom{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}}\right\} \text{4 rows}$$

$\mathbf{v} \in \mathbb{R}^{\mathbf{1\times 3}}$

$\mathbf{u} \in \mathbb{R}^{\mathbf{3\times 1}}$

$\mathbf{A} \in \mathbb{R}^{\mathbf{4\times 3}}$

Matrix Multiplication

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 5 & 7 \end{bmatrix} \rightarrow \quad \mathbf{C} = [\quad] \quad ?$$

4×3 $\qquad$ 3×2

Dimensions much match!

What is the dimension of C? $(4 \times 3)(3 \times 2) \rightarrow 4 \times 2$

Matrix Multiplication

$1 \times 1 + 2 \times 2 + 3 \times 5$

$$\mathbf{A} = \begin{bmatrix} \cancel{1} & \cancel{2} & \cancel{3} \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \quad \mathbf{B} = \begin{bmatrix} \cancel{1} & 2 \\ \cancel{2} & 3 \\ \cancel{5} & 7 \end{bmatrix} \Rightarrow \quad \mathbf{C} = \begin{bmatrix} \mathbf{20} & \\ & \\ & \end{bmatrix}$$

4×3 $\qquad$ 3×2

$1 \times 2 + 2 \times 3 + 3 \times 7$

$$\mathbf{A} = \begin{bmatrix} \cancel{1} & \cancel{2} & \cancel{3} \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \quad \mathbf{B} = \begin{bmatrix} 1 & \cancel{2} \\ 2 & \cancel{3} \\ 5 & \cancel{7} \end{bmatrix} \Rightarrow \quad \mathbf{C} = \begin{bmatrix} 20 & \mathbf{29} \\ & \\ & \end{bmatrix}$$

4×3 $\qquad$ 3×2

Matrix Multiplication

$$0 \times 1 + 5 \times 2 + 1 \times 5$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 5 & 7 \end{bmatrix} \Rightarrow \mathbf{C} = \begin{bmatrix} 20 & 29 \\ \mathbf{15} & \\ & \\ & \end{bmatrix}$$

$$4 \times 3 \qquad 3 \times 2$$

$$0 \times 2 + 5 \times 3 + 1 \times 7$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 5 & 7 \end{bmatrix} \Rightarrow \mathbf{C} = \begin{bmatrix} 20 & 29 \\ 15 & \mathbf{22} \\ & \\ & \end{bmatrix}$$

$$4 \times 3 \qquad 3 \times 2$$

Basics – Linear Algebra

Matrix Multiplication

$$2 \times 1 + 3 \times 2 + 7 \times 5$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 5 & 7 \end{bmatrix} \Rightarrow \mathbf{C} = \begin{bmatrix} 20 & 29 \\ 15 & 22 \\ \mathbf{43} & \\ & \end{bmatrix}$$

$$4 \times 3 \qquad 3 \times 2$$

$$2 \times 2 + 3 \times 3 + 7 \times 7$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 5 & 7 \end{bmatrix} \Rightarrow \mathbf{C} = \begin{bmatrix} 20 & 29 \\ 15 & 22 \\ 43 & \mathbf{62} \\ & \end{bmatrix}$$

$$4 \times 3 \qquad 3 \times 2$$

Matrix Multiplication

$$3 \times 1 + 2 \times 9 + 5 \times 8$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 5 & 7 \end{bmatrix} \Rightarrow \mathbf{C} = \begin{bmatrix} 20 & 29 \\ 15 & 22 \\ 43 & 62 \\ \mathbf{61} & \end{bmatrix}$$

4×3 $\quad$ 3×2

$$3 \times 2 + 3 \times 9 + 7 \times 8$$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 2 & 3 & 7 \\ 3 & 9 & 8 \end{bmatrix} \times \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 5 & 7 \end{bmatrix} \Rightarrow \mathbf{C} = \begin{bmatrix} 20 & 29 \\ 15 & 22 \\ 43 & 62 \\ 61 & \mathbf{89} \end{bmatrix}$$

4×3 $\quad$ 3×2

Basics – Linear Algebra

- $\|\mathbf{u}-\mathbf{v}\|^p$
- Function from a vector space to a single positive real value: $f: \mathbb{R}^d \rightarrow \mathbb{R}$
- Distance between **u** and **v**

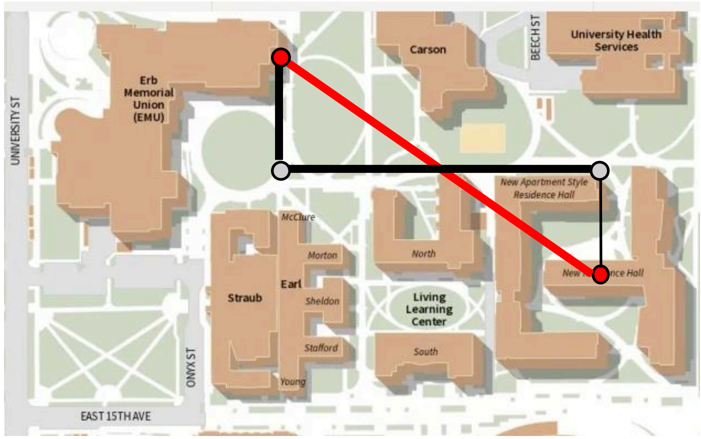
$$\|\mathbf{u}-\mathbf{v}\|^p = \left(\sum_{i=1}^{d} |\mathbf{u}_i - \mathbf{v}_i|^p\right)^{\frac{1}{p}}$$

- Examples:

(1) Manhattan distance (L_1): $\|\mathbf{u}-\mathbf{v}\|^1 = \left(\sum_{i=1}^{d} |\boldsymbol{u}_i - \boldsymbol{v}_i|\right)$

(2) Euclidean distance (L_2): $\|\mathbf{v}\|^2 = \left(\sum_{i=1}^{d} |\mathbf{u}_i - \mathbf{v}_i|^2\right)^{\frac{1}{2}}$

  - **How about L_0?**



1 Basics

$$
\begin{aligned}
(\mathbf{AB})^{-1} &= \mathbf{B}^{-1}\mathbf{A}^{-1} && (1)\\
(\mathbf{ABC}...)^{-1} &= ...\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1} && (2)\\
(\mathbf{A}^T)^{-1} &= (\mathbf{A}^{-1})^T && (3)\\
(\mathbf{A}+\mathbf{B})^T &= \mathbf{A}^T+\mathbf{B}^T && (4)\\
(\mathbf{AB})^T &= \mathbf{B}^T\mathbf{A}^T && (5)\\
(\mathbf{ABC}...)^T &= ...\mathbf{C}^T\mathbf{B}^T\mathbf{A}^T && (6)\\
(\mathbf{A}^H)^{-1} &= (\mathbf{A}^{-1})^H && (7)\\
(\mathbf{A}+\mathbf{B})^H &= \mathbf{A}^H+\mathbf{B}^H && (8)\\
(\mathbf{AB})^H &= \mathbf{B}^H\mathbf{A}^H && (9)\\
(\mathbf{ABC}...)^H &= ...\mathbf{C}^H\mathbf{B}^H\mathbf{A}^H && (10)
\end{aligned}
$$

Matrix Codebook

https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf

The Matrix Cookbook

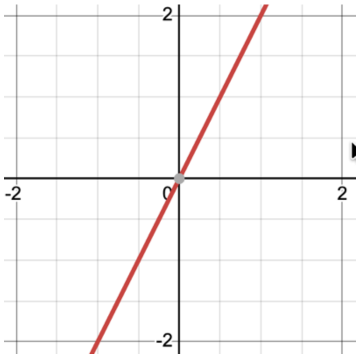
[http://matrixcookbook.com]

Kaare Brandt Petersen
Michael Syskind Pedersen

VERSION: NOVEMBER 15, 2012

Basics – Derivative and Gradient

$$y = 2x, \qquad \frac{dy}{dx} = 2$$



$$y = 2x_1 + 3x_2, \qquad \frac{\partial y}{\partial x_1} = 2, \frac{\partial y}{\partial x_2} = 3 \qquad \mathbf{a} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$y = \mathbf{a}^{\mathrm{T}}\mathbf{x}$$

$$\frac{\partial y}{\partial \boldsymbol{x}} = \boldsymbol{a}$$

$$y_1 = 2x_1 + 3x_2, \qquad \frac{\partial y_1}{\partial x_1} = 2, \frac{\partial y_2}{\partial x_2} = 3 \qquad \mathbf{a} = \begin{bmatrix} 2 & 4 \\ 3 & 3 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

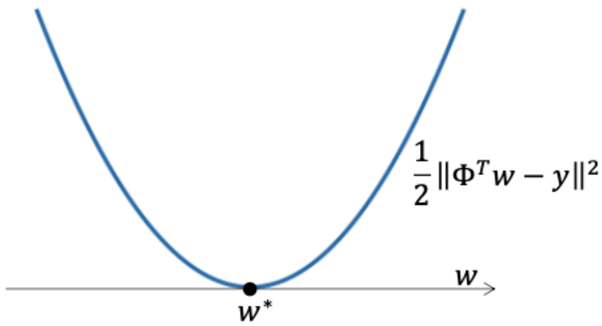
$$y_2 = 4x_1 + 5x_2, \qquad \frac{\partial y_2}{\partial x_1} = 4, \frac{\partial y_2}{\partial x_2} = 3 \qquad \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \boldsymbol{a}$$

$$\boldsymbol{y} = \mathbf{a}^{\mathrm{T}}\mathbf{x}$$

How to use the book?

Goal: $\underset{w}{\text{argmin}}\ \frac{1}{2}\left\|\Phi^{\mathrm{T}}w - y\right\|^2$

$$\mathcal{L}(w) = \frac{1}{2}\left\|\Phi^{\mathrm{T}}w - y\right\|^2 = \frac{1}{2}\left(\Phi^{\mathrm{T}}w - y\right)^T\left(\Phi^{\mathrm{T}}w - y\right)$$
$$= \frac{1}{2}(w^T\Phi - y^T)(\Phi^T w - y)$$
$$= \frac{1}{2}w^T\Phi\Phi^T w - w^T\Phi y + \frac{1}{2}y^T y$$



Calculate the gradient of loss w.r.t. the model parameters?

Basics – Linear Algebra

How to use the book?

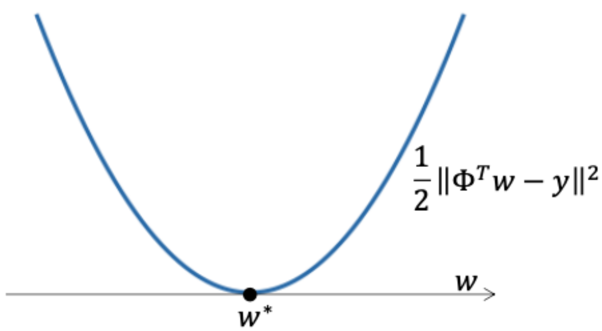
Goal: $\underset{w}{\operatorname{argmin}} \frac{1}{2}\left\|\Phi^{\mathrm{T}}w - y\right\|^2$

$$\mathcal{L}(w) = \frac{1}{2}\left\|\Phi^{\mathrm{T}}w - y\right\|^2 = \frac{1}{2}\left(\Phi^{\mathrm{T}}w - y\right)^T\left(\Phi^{\mathrm{T}}w - y\right)$$

$$= \frac{1}{2}\left(w^T\Phi - y^T\right)\left(\Phi^T w - y\right)$$

$$= \frac{1}{2}w^T\Phi\Phi^T w - w^T\Phi y + \frac{1}{2}y^T y$$

$$\nabla_w \mathcal{L} = \Phi\Phi^T w - \Phi y$$



Gradient descent solution:

$$w^{(t+1)} = w^{(t)} - \epsilon\nabla_w \mathcal{L}(w^{(t)})$$

2.4 Derivatives of Matrices, Vectors and Scalar Forms

2.4.1 First Order

$$\frac{\partial \mathbf{x}^T\mathbf{a}}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}} = \mathbf{a} \tag{69}$$

$$\frac{\partial \mathbf{a}^T\mathbf{X}\mathbf{b}}{\partial \mathbf{X}} = \mathbf{a}\mathbf{b}^T \tag{70}$$

$$\frac{\partial \mathbf{x}^T\mathbf{B}\mathbf{x}}{\partial \mathbf{x}} = (\mathbf{B}+\mathbf{B}^T)\mathbf{x} \tag{81}$$

$$\frac{\partial \mathbf{b}^T\mathbf{X}^T\mathbf{D}\mathbf{X}\mathbf{c}}{\partial \mathbf{X}} = \mathbf{D}^T\mathbf{X}\mathbf{b}\mathbf{c}^T + \mathbf{D}\mathbf{X}\mathbf{c}\mathbf{b}^T \tag{82}$$

$$\frac{\partial}{\partial \mathbf{X}}(\mathbf{X}\mathbf{b}+\mathbf{c})^T\mathbf{D}(\mathbf{X}\mathbf{b}+\mathbf{c}) = (\mathbf{D}+\mathbf{D}^T)(\mathbf{X}\mathbf{b}+\mathbf{c})\mathbf{b}^T \tag{83}$$

- **Probability → *from model to data***

A fair coin
P(Head) = 0.5

What is the probability of observing 7 heads in 10 tosses?

- **Statistics → *from data to model*** ***(machine learning as well)***

Data: 10 tosses → 7 heads

Question: Is the coin fair? What is p?

Basics – Probability

- **Sample Space: The set of all possible outcomes**
- **Event: A subset of the sample space**
- **Probability: under certain situation, how much likelihood of event**

Space {1, 2, 3, 4, 5, 6}

"Rolling an even number" = {2, 4, 6}

Basics – Probability

1) Possible values for probabilities range from 0 to 1

0 = impossible event

1 = certain event

2) The sum of all the probabilities for all possible outcomes is equal to 1.

Note the connection to the complement rule.

3) Addition Rule - the probability that one or both events occur

mutually exclusive events: P(A or B) = P(A) + P(B)

not mutually exclusive events: P(A or B) = P(A) + P(B) - P(A and B)

4) Multiplication Rule - the probability that **both** events occur together

independent events: P(A and B) = P(A) * P(B)

P(A and B) = P(A) * P(B|A)

5) Conditional Probability - the probability of an event happening **given** that another event has already happened

P(A|B) = P(A and B) / P(B)

*Note the line | means "given" while the slash / means divide

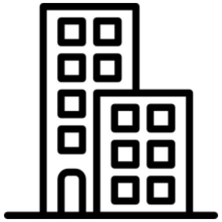
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' Rule

Basics – Probability

P(B=W) = 0.3

P(ND|W) = 0.6



P(B=G) = 0.5

P(ND|G) = 0.2

P(B=S) = 0.2

P(ND|S) = 0.05

Basics – Probability

P(B=W) = 0.3

P(B=G) = 0.5

P(B=S) = 0.2

P(ND|W) = 0.4

P(ND|G) = 0.8

P(ND|S) = 0.95

After one earthquake, the building is not collapsed

$$P(T|ND) = \frac{P(ND|T)P(T)}{P(ND)}$$

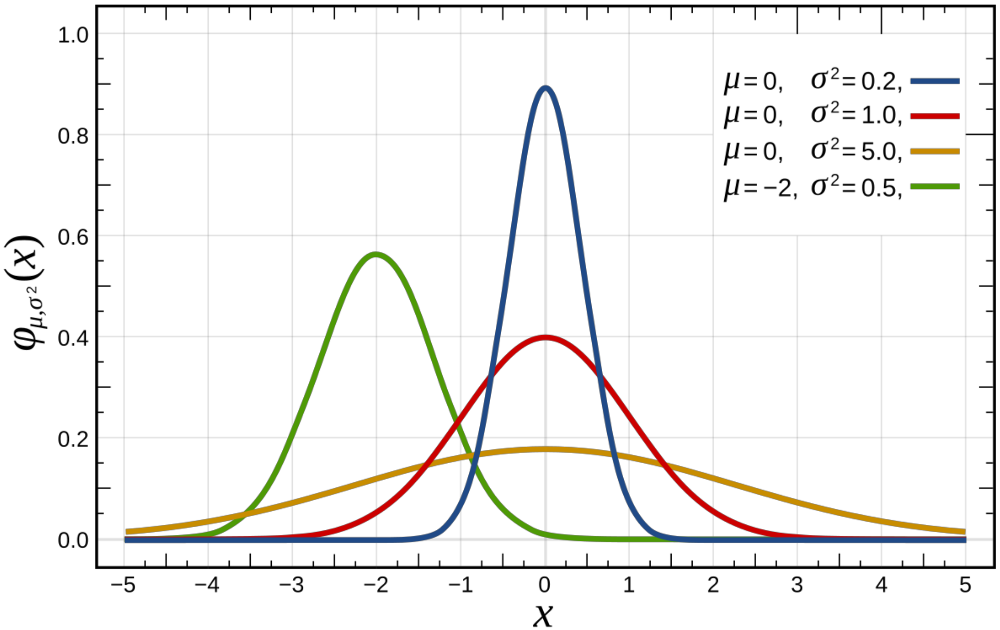
$$P(ND) = \sum_T P(ND|T)P(T)$$

$$P(G|ND) = \frac{0.8 * 0.5}{0.71} = 0.56 \qquad P(S|ND) = \frac{0.95 * 0.2}{0.71} = 0.27$$

$$P(W|ND) = \frac{P(ND|W)P(W)}{P(ND)} = \frac{0.4 * 0.3}{0.71} = 0.17$$

$$P(ND) = \sum_T P(ND|T)P(T)$$
$$= 0.3 * 0.4 + 0.5 * 0.8 + 0.2 * 0.95 = 0.71$$

Basics – Probability Density Function – Distribution

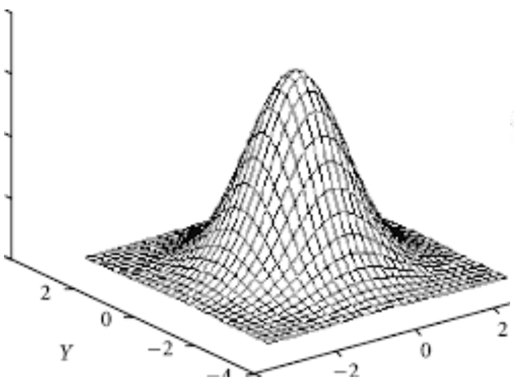


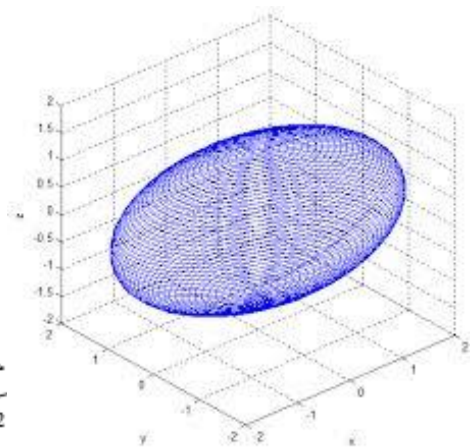
1-D Probability Density Function

2-D Probability Density Function

3-D Probability Density Function

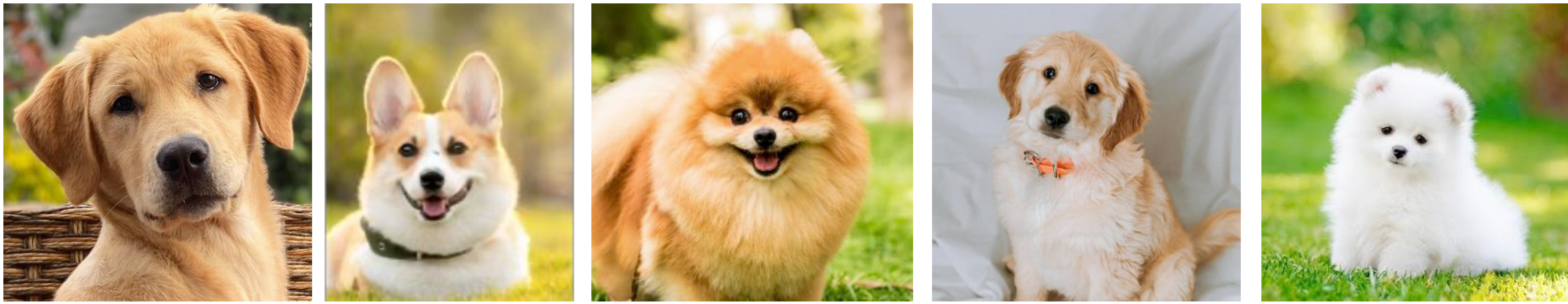
N-D Probability Density Function





Basics – High Dimensional Random Variable

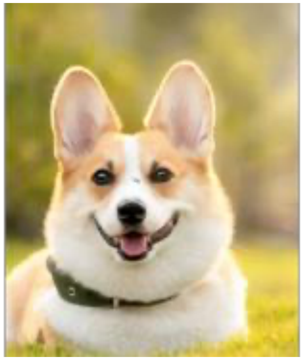
Dog



Cat

Basics – High Dimensional Random Variable

Dog – P(Dog)

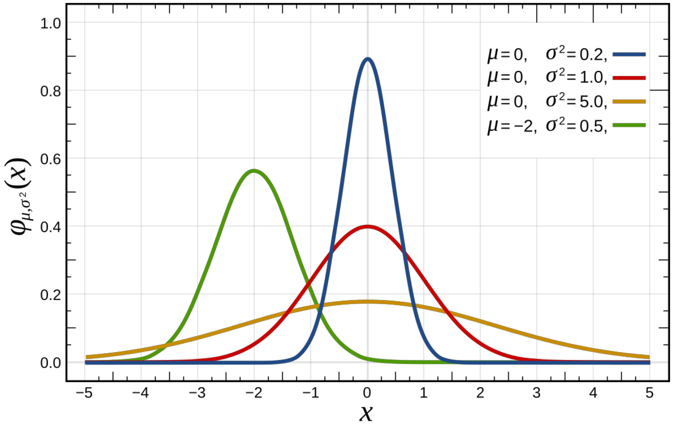


Cat – P(Cat)

1. **There is no concrete image/shape of the dog, everyone can come up with one of your own choice**

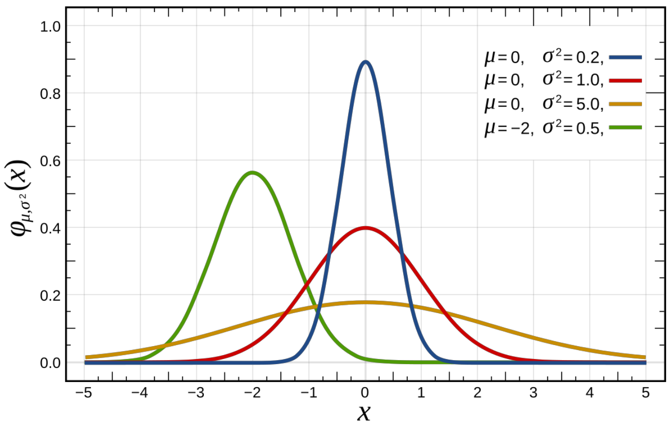
2. **But somehow dog and cat image distributions are different**

When you draw an image, you are actually sampling from a probability distribution!



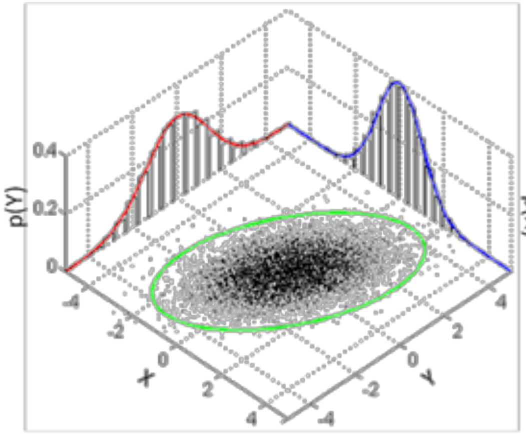
Basics – Data Distribution

1D Gaussian Distribution



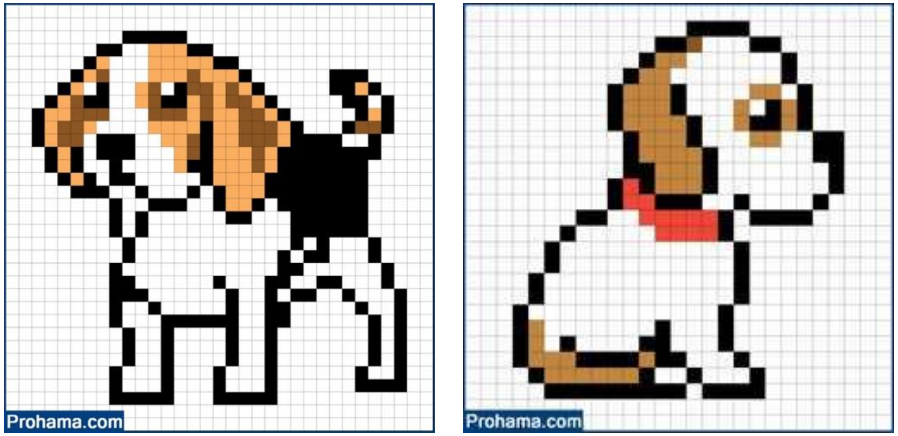
$\mathbb{R}$

2D Gaussian Distribution



$\mathbb{R}^2$

$\mathbb{R}^{256\times256}$



$\mathbb{R}^{256\times256}$