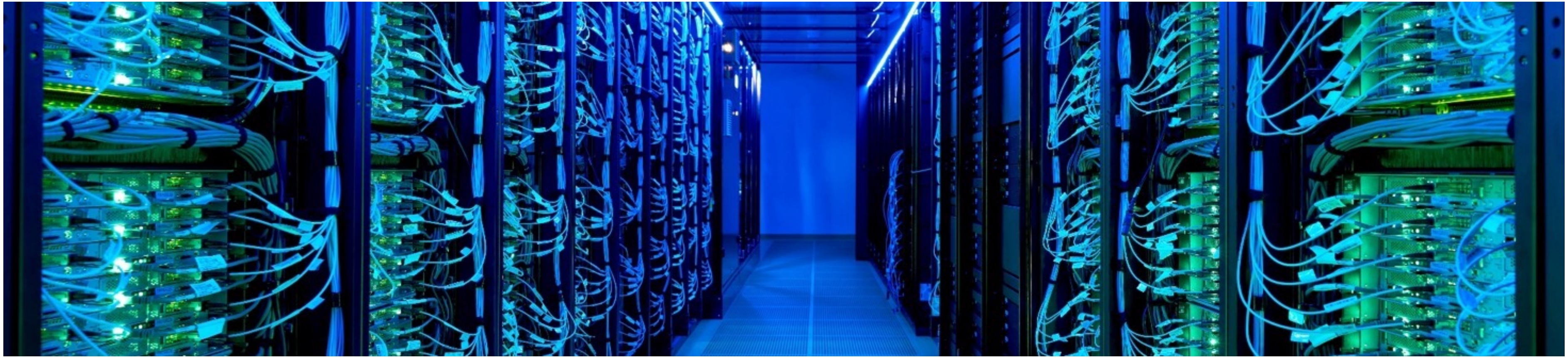
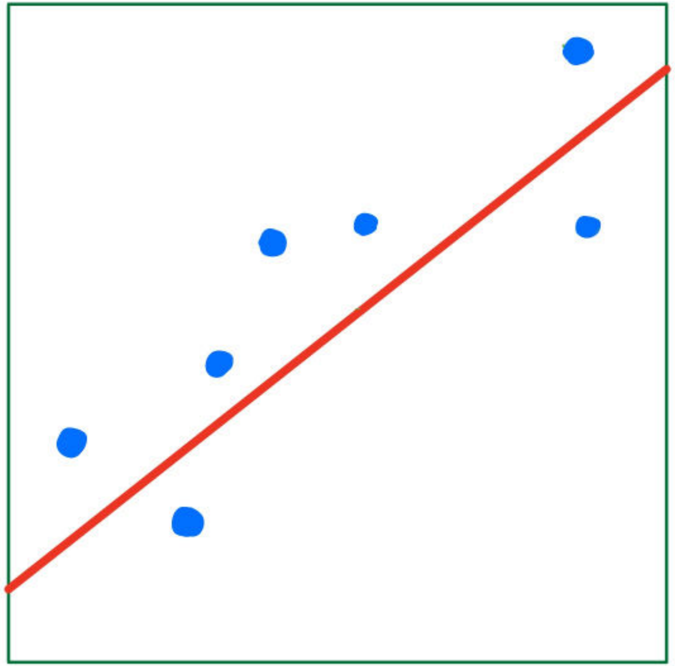
# How much mathematical modeling is in AI?

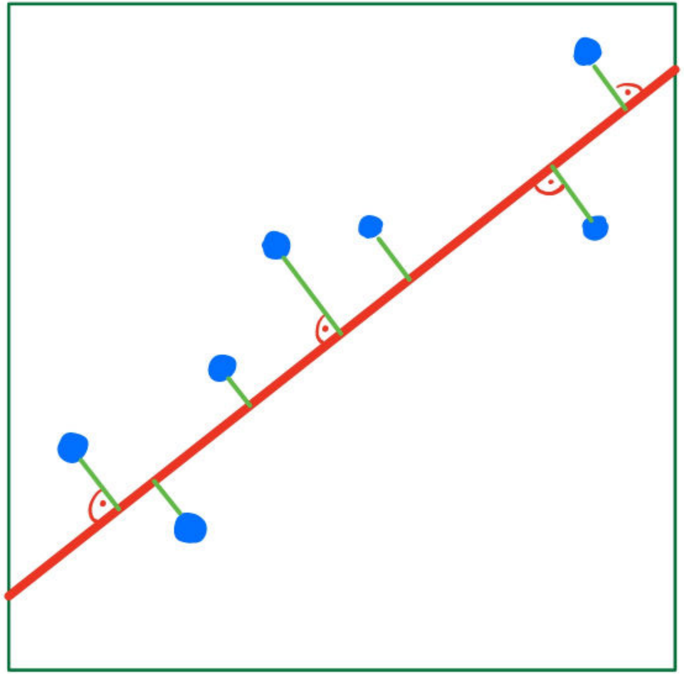
Martin Frank, Sarah Schönbrodt



# Linear Regression

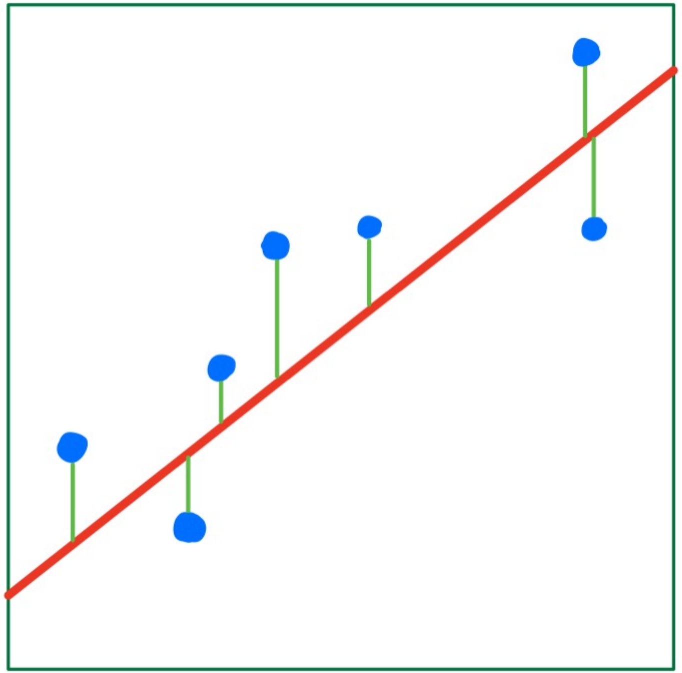


# Linear Regression



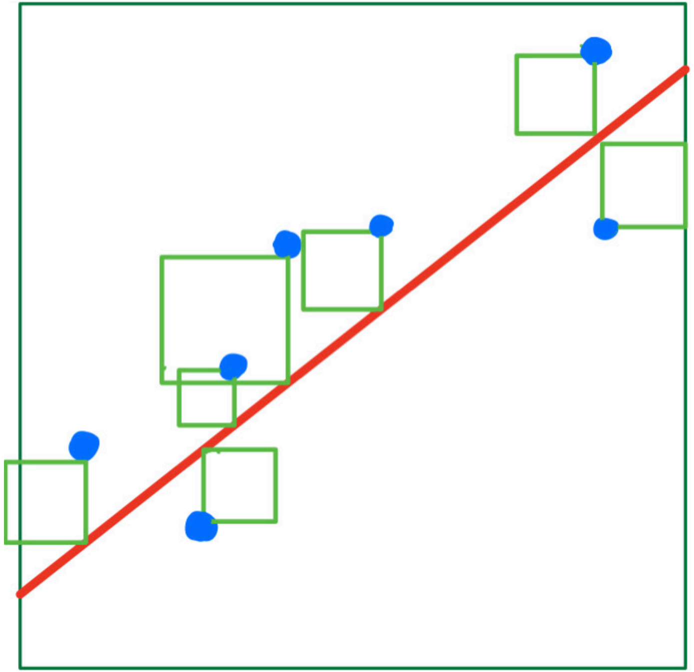
sum of minimal distances

# Linear Regression



sum of y-distances

# Linear Regression

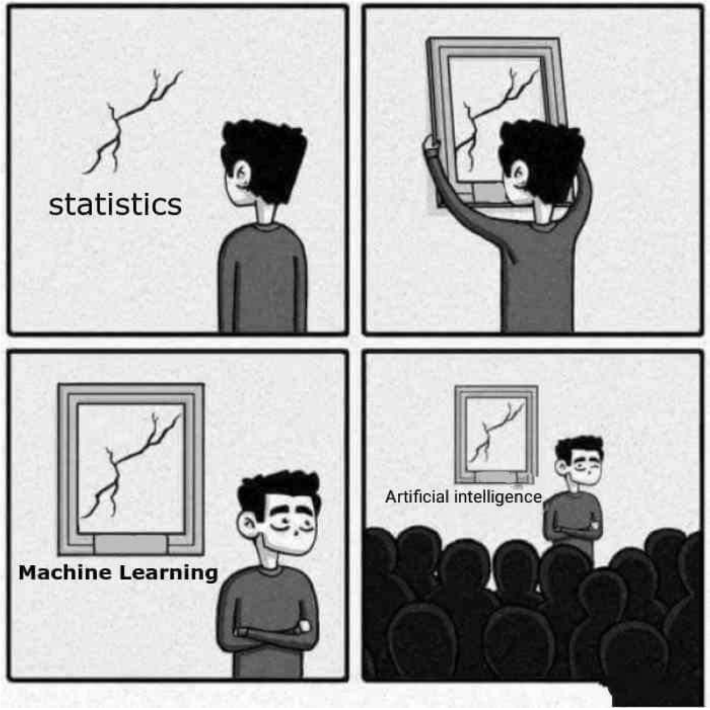


sum of squared y-distances

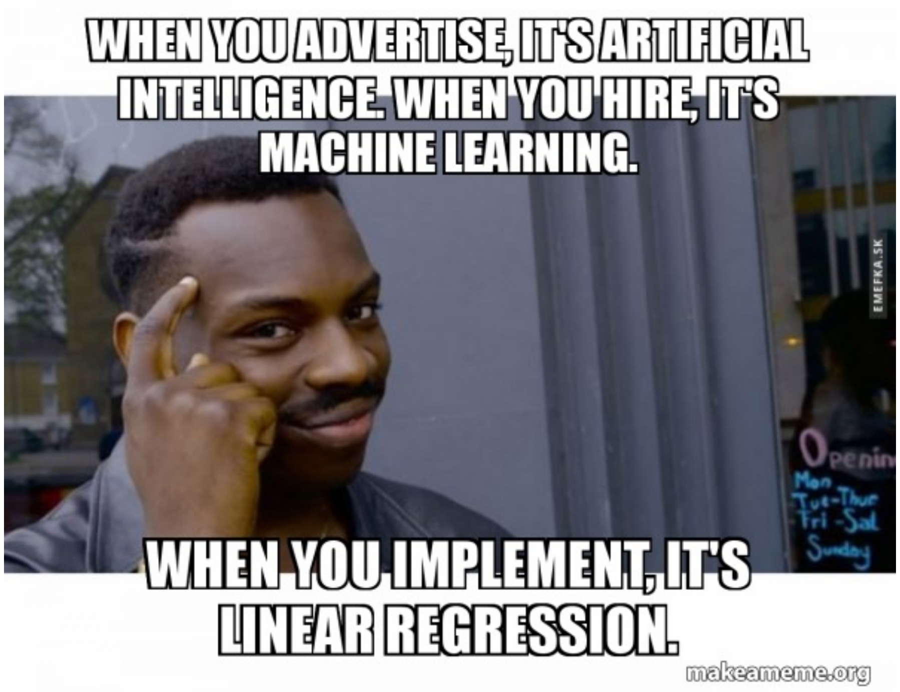
# Take-Home Message

DS/ML/AI are mathematical modeling challenges $\wedge$ Mathematical modeling belongs into Mathematics teaching $\Longrightarrow$ AI belongs into the Mathematics curriculum in schools

# AI is a Hype




Joe Davison, https://towardsdatascience.com




Roman Orac, https://romanorac.medium.com/

# AI is real

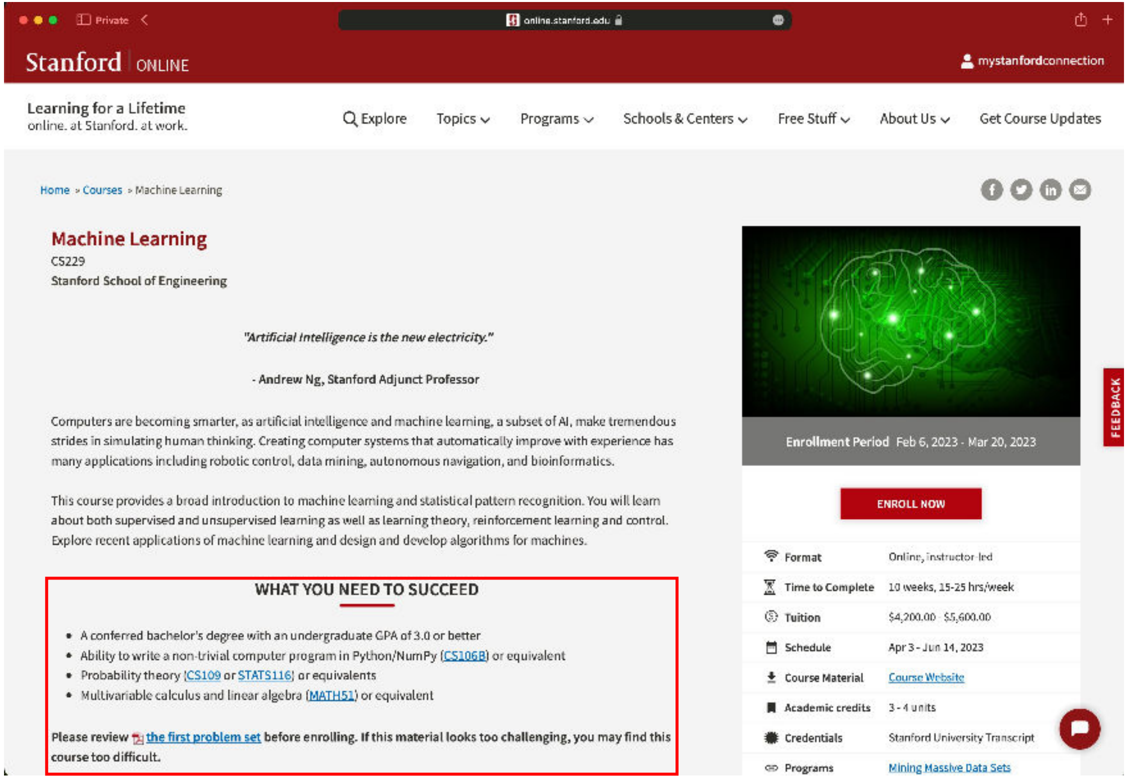
Known in the public:

- ChatGPT → everybody is talking about it
- DALL-E/Midjourney → everybody should be talking about it more

Scientific breakthroughs

- AlphaFold → protein folding
- Natural language translation used for biological “languages”
- FourCastNet → weather forecast

# How much Mathematics is in AI?



Private

online.stanford.edu

Stanford ONLINE

mystanfordconnection

Learning for a Lifetime
online. at Stanford. at work.

Explore | Topics | Programs | Schools & Centers | Free Stuff | About Us | Get Course Updates

Home » Courses » Machine Learning

## Machine Learning

CS229

Stanford School of Engineering

*"Artificial Intelligence is the new electricity."*

- Andrew Ng, Stanford Adjunct Professor

Computers are becoming smarter, as artificial intelligence and machine learning, a subset of AI, make tremendous strides in simulating human thinking. Creating computer systems that automatically improve with experience has many applications including robotic control, data mining, autonomous navigation, and bioinformatics.

This course provides a broad introduction to machine learning and statistical pattern recognition. You will learn about both supervised and unsupervised learning as well as learning theory, reinforcement learning and control. Explore recent applications of machine learning and design and develop algorithms for machines.

### WHAT YOU NEED TO SUCCEED

- A conferred bachelor's degree with an undergraduate GPA of 3.0 or better
- Ability to write a non-trivial computer program in Python/NumPy (CS106B) or equivalent
- Probability theory (CS109 or STATS116) or equivalents
- Multivariable calculus and linear algebra (MATH51) or equivalent

Please review the first problem set before enrolling. If this material looks too challenging, you may find this course too difficult.

Enrollment Period Feb 6, 2023 - Mar 20, 2023

ENROLL NOW

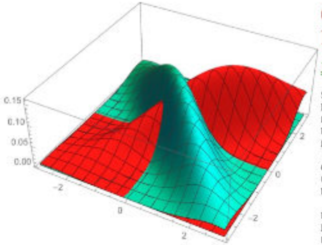
| | |
|---|---|
| Format | Online, instructor-led |
| Time to Complete | 10 weeks, 15-25 hrs/week |
| Tuition | \$4,200.00 - \$5,600.00 |
| Schedule | Apr 3 - Jun 14, 2023 |
| Course Material | Course Website |
| Academic credits | 3 - 4 units |
| Credentials | Stanford University Transcript |
| Programs | Mining Massive Data Sets |

FEEDBACK

https://online.stanford.edu/courses/cs229-machine-learning

# How much Mathematics is in AI?

Karlsruhe Institute of Technology



## CS 189/289A
## Introduction to Machine Learning

Jonathan Shewchuk

Spring 2023
Mondays and Wednesdays, 6:30–8:00 pm
Hearst Field Annex A1
Begins Wednesday, January 18
Discussion sections begin Tuesday, January 24

Contact:
Use Ed Discussion for public and private questions that can be viewed by all the TAs. I check Ed Discussion far more often and reliably than email.
For very personal issues, send email to jrs@berkeley.edu.

My office hours:
Mondays, 5:10–6:00 pm
Fridays, 5:40–6:30 pm
and by appointment.
(I'm usually free after the lectures too.)

This class introduces algorithms for *learning*, which constitute an important part of artificial intelligence.

Topics include

- classification: perceptrons, support vector machines (SVMs), Gaussian discriminant analysis (including linear discriminant analysis, LDA, and quadratic discriminant analysis, QDA), logistic regression, decision trees, neural networks, convolutional neural networks, boosting, nearest neighbor search;
- regression: least-squares linear regression, logistic regression, polynomial regression, ridge regression, Lasso;
- density estimation: maximum likelihood estimation (MLE);
- dimensionality reduction: principal components analysis (PCA), random projection; and
- clustering: *k*-means clustering, hierarchical clustering.

## Useful Links

- Access the CS 189/289A Ed Discussion forum. If you haven't already been added to the class, use this invitation link.
- Submit your assignments at the CS 189/289A Gradescope. If you need the entry code, find it on Ed Discussion in the post entitled "Welcome to CS 189!"
- If you want an instructional account, you can get one online. Go to the same link if you forget your password or account name.
- Check out this Machine Learning Visualizer by our former TA Sagnik Bhattacharya and his teammates Colin Zhou, Komila Khamidova, and Aaron Sun. It's a great way to build intuition for what decision boundaries different classification algorithms find.

## Prerequisites

- Math 53 (or another vector calculus course),
- Math 54, Math 110, or EE 16A+16B (or another linear algebra course),
- CS 70, EECS 126, or Stat 134 (or another probability course).
- Enough programming experience to be able to debug complicated programs without much help. (Unlike in a lower-division programming course, the Teaching Assistants are under no obligation to look at your code.)

You should take these prerequisites quite seriously. If you don't have a solid intuitive understanding of linear algebra, probability, and gradients, as well as substantial programming experience with some attention to data structures, I strongly recommend not taking CS 189. However, the prerequisites are not *formally* enforced—rather, they're enforced by the fact that you won't understand the class without them.

If you want to brush up on prerequisite material:

- Here's a short summary of math for machine learning written by our former TA Garrett Thomas.
- Stanford's machine learning class provides additional reviews of linear algebra and probability theory.
- There's a fantastic collection of linear algebra visualizations on YouTube by 3Blue1Brown starting with this playlist, *The Essence of Linear Algebra*. I highly recommend them, even if you think you already understand linear algebra. It's not enough to know how to work with matrix algebra equations; it's equally important to have a geometric intuition for what it all means.
- To learn matrix calculus (which will rear its head first in Homework 2), check out the first two chapters of The Matrix Cookbook.
- Another locally written review of linear algebra appears in this book by Prof. Laurent El Ghaoui.
- An alternative guide to CS 189 material (if you're looking for a second set of lecture notes besides mine), written by our former TAs Soroush Nasiriany and Garrett Thomas, is available at this link. I recommend reading my notes first, but reading the same material presented a different way can help you firm up your understanding.

https://people.eecs.berkeley.edu/~jrs/189/

# How much Mathematics is in AI?

Private people.eecs.berkeley.edu

This class introduces algorithms for *learning*, which constitute an important part of artificial intelligence.

Topics include

- classification: perceptrons, support vector machines (SVMs), Gaussian discriminant analysis (including linear discriminant analysis, LDA, and quadratic discriminant analysis, QDA), logistic regression, decision trees, neural networks, convolutional neural networks, boosting, nearest neighbor search;
- regression: least-squares linear regression, logistic regression, polynomial regression, ridge regression, Lasso;
- density estimation: maximum likelihood estimation (MLE);
- dimensionality reduction: principal components analysis (PCA), random projection; and
- clustering: $k$-means clustering, hierarchical clustering.

## Useful Links

- Access the CS 189/289A Ed Discussion forum. If you haven't already been added to the class, use this invitation link.
- Submit your assignments at the CS 189/289A Gradescope. If you need the entry code, find it on Ed Discussion in the post entitled "Welcome to CS 189!"
- If you want an instructional account, you can get one online. Go to the same link if you forget your password or account name.
- Check out this Machine Learning Visualizer by our former TA Sagnik Bhattacharya and his teammates Colin Zhou, Komila Khamidova, and Aaron Sun. It's a great way to build intuition for what decision boundaries different classification algorithms find.

## Prerequisites

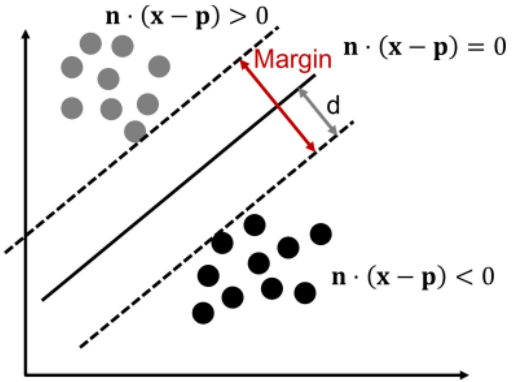
- Math 53 (or another vector calculus course),
- Math 54, Math 110, or EE 16A+16B (or another linear algebra course),
- CS 70, EECS 126, or Stat 134 (or another probability course).
- Enough programming experience to be able to debug complicated programs without much help. (Unlike in a lower-division programming course, the Teaching Assistants are under no obligation to look at your code.)

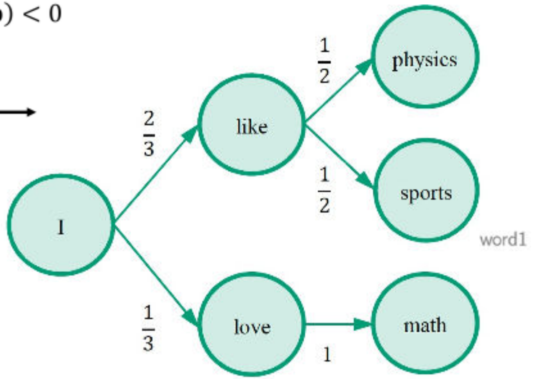
You should take these prerequisites quite seriously. If you don't have a solid intuitive understanding of linear algebra, probability, and gradients, as well as substantial programming experience with some attention to data structures, I strongly recommend not taking CS 189. However, the prerequisites are not *formally* enforced—rather, they're enforced by the fact that you won't understand the class without them.

https://people.eecs.berkeley.edu/~jrs/189/

# Mathematics of AI is accessible

- Existing learning material:
  - Support Vector Machines
  - Word prediction
  - K-Nearest Neighbors, k-Means
  - Latent factor models
  - Statistical evaluation
  - Inverse problems
  - Optimization
- In preparation:
  - Neural networks as concatenations of elementary functions
  - Data cleaning




**How to actually do it → Sarah**

# Take-Home Message

DS/ML/AI are mathematical modeling challenges

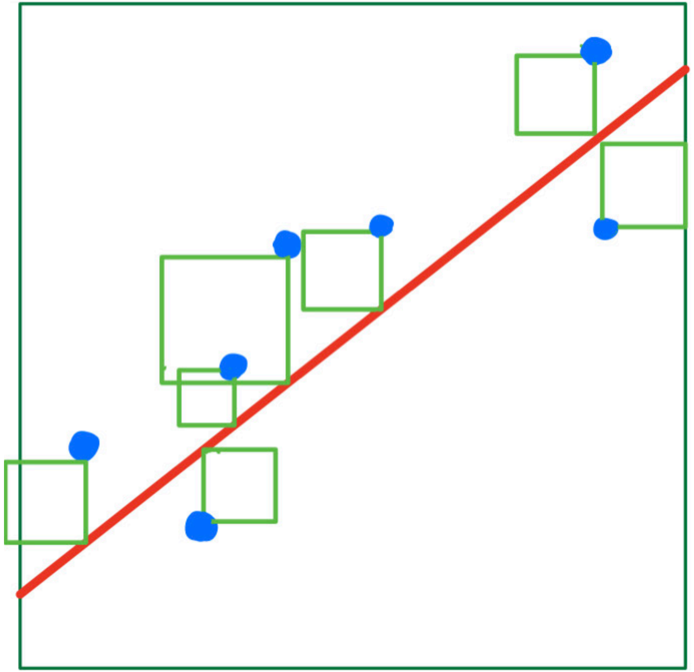
$\wedge$

Mathematical modeling belongs into Mathematics teaching

$\Longrightarrow$

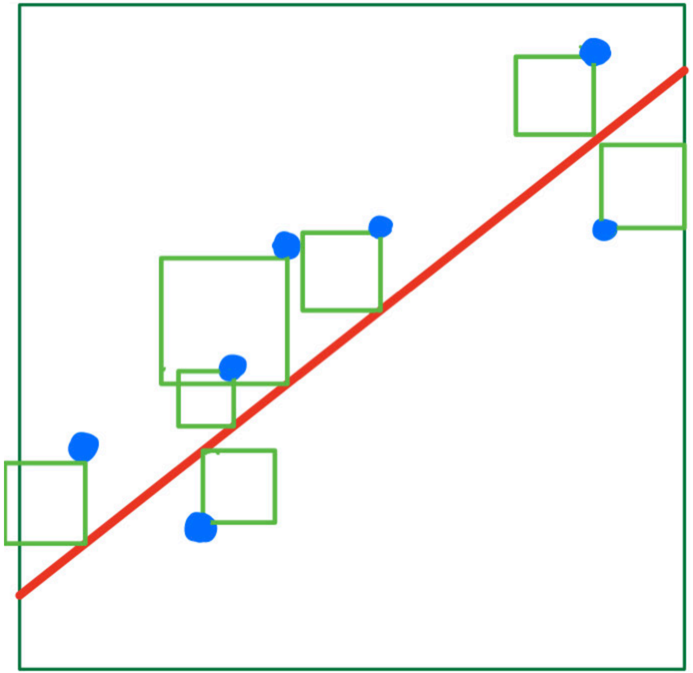
AI belongs into the Mathematics curriculum in schools

# Linear Regression



sum of squared y-distances

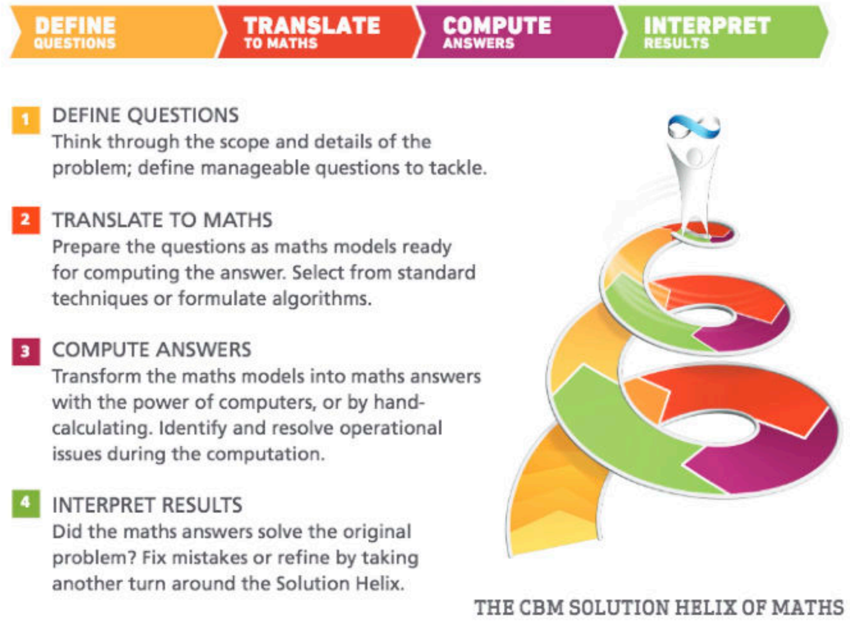
# Mathematical Modeling belongs into Mathematics



- The reasons for using the least squares functional are mathematical: Optimization & Statistics

- Modeling is the process of translating a real-world problem with a certain purpose into a mathematically treatable problem/description; it is a subtle interplay of what is necessary to address a problem, and what is mathematically doable

# Mathematical Modeling belongs into Mathematics

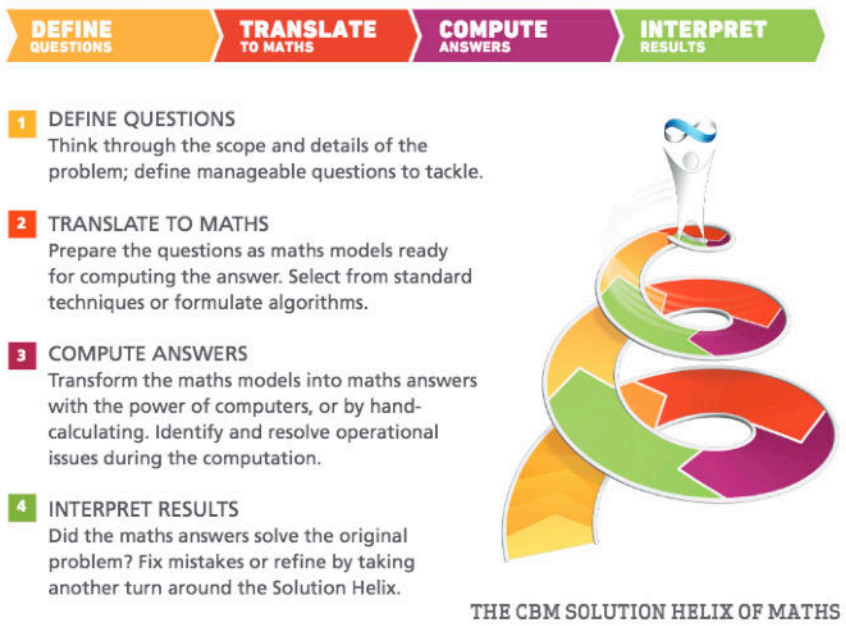



Source: computerbasedmath.org

- The reasons for using the least squares functional are mathematical: Optimization & Statistics

- Modeling is the process of translating a real-world problem with a certain purpose into a mathematically treatable problem/description; it is a subtle interplay of what is necessary to address a problem, and what is mathematically doable

# Mathematics Education needs to be adapted




Source: computerbasedmath.org

- Substitute mechanical task of computing on paper by computer use
- Modeling requires active, creative use of Mathematics
- Mathematical modeling (also AI and DS) are inherently interdisciplinary

# Quod Erat Demonstrandum

- DS/ML/AI are mathematical modeling challenges

$\wedge$

- Mathematical modeling belongs into Mathematics teaching

$\Longrightarrow$

- AI belongs into the Mathematics curriculum in schools

$\wedge$

- It can be done!
  → Sarah

# Part II – It can be done!

# Education Lab CAMMP

**Computational And Mathematical Modeling Program**

**Goal**

Allows students to solve

**authentic, real** and **relevant problems**

with the help of **mathematical modeling** and **computers**.

- teaching and learning material
- modeling activities for students and teachers
- AI and also non AI related topics

# Teaching formats

- **pre-structured activities** on problems related to everyday life
  - modeling / project days or lesson series
  - **goal:** teach mathematical background of selected AI method $\Rightarrow$ sophisticated models
- **open activities** on current problems of companies / institutes
  - modeling / project weeks
  - **goal:** allow maximum creativity when modeling (data-intensive) problems $\Rightarrow$ (possibly) simpler models

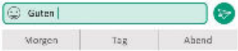
**Challenge:** Balance guidance vs. mathematical creativity

# Our AI workshops

### Security in social networks

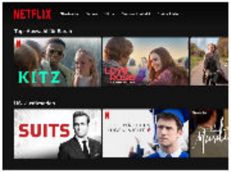
Sube (2019)

### Word predicitions




Hofmann (2022)

### Recommender systems




Schönbrodt et al. (2021)

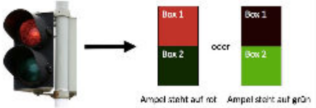
**Problems**

### Activity classification on smartphones

phyphox
physical phone experiments

Running? Walking? Climbing stairs?

### (Image) Classification




Schönbrodt (2019); Schönbrodt et al. (2021)

# Example: Recommender systems

## Netflix Prize

- task: improve Netflix recommendation system by at least 10 %
- main prize: 1,000,000 $
- winner: Team *BellKor's Pragmatic Chaos* (2009)
- Netflix published a huge dataset in October 2006

**The problem, some learning material and experiences with students (pre-structured and open activities)**

# Example: Recommender systems

## Dataset of the Netflix Prize

- 17,770 movies
- 480,189 user
- 100,480,507 ratings (from 1 to 5) ★★★★★

$$\begin{array}{cc} & \begin{array}{ccccccc} {}_1 & {}_2 & {}_3 & {}_4 & {}_5 & {}_6 & \cdots \end{array} \\ \begin{array}{c} {}_1 \\ {}_2 \\ {}_3 \\ {}_4 \\ {}_5 \\ \vdots \end{array} & \begin{pmatrix} 3 & ? & 1 & ? & 1 & 4 & \cdots \\ ? & 2 & 4 & 1 & 3 & 1 & \cdots \\ 3 & 1 & ? & 3 & ? & ? & \cdots \\ 4 & 3 & ? & 4 & 4 & ? & \cdots \\ 4 & ? & ? & 4 & ? & 5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}$$

# Example: Recommender systems

## Dataset of the Netflix Prize

- 17,770 movies 🎞
- 480,189 user 👤
- 100,480,507 ratings (from 1 to 5) ★★★★★

$$
\begin{array}{c}
 \\
\text{👤}_1 \\
\text{👤}_2 \\
\text{👤}_3 \\
\text{👤}_4 \\
\text{👤}_5 \\
\vdots
\end{array}
\begin{array}{c}
\begin{array}{ccccccc}
\text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \text{🎞}_4 & \text{🎞}_5 & \text{🎞}_6 & \cdots
\end{array} \\
\begin{pmatrix}
3 & \mathbf{2} & 1 & \mathbf{5} & 1 & 4 & \cdots \\
\mathbf{1} & 2 & 4 & 1 & 3 & 1 & \cdots \\
3 & 1 & \mathbf{3} & 3 & \mathbf{2} & \mathbf{1} & \cdots \\
4 & 3 & \mathbf{1} & 4 & 4 & \mathbf{2} & \cdots \\
4 & \mathbf{2} & \mathbf{2} & 4 & \mathbf{3} & 5 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\end{array}
$$

**Task:** Predict ratings in the best possible way!

# From research into high-schools I

$$
R = \begin{array}{c} \\ {}_1 \\ {}_2 \\ {}_3 \\ {}_4 \\ {}_5 \end{array}
\begin{array}{c}
\begin{array}{cccccc} {}_1 & {}_2 & {}_3 & {}_4 & {}_5 & {}_6 \end{array} \\
\begin{pmatrix}
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & 5 & 4 & 3 & 5 \\
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & \color{red}{?} & 4 & 3 & 5 \\
1 & \color{red}{?} & 4 & 1 & 1 & 2
\end{pmatrix}
\end{array}
$$

**Basic idea:**

Model hidden movie features & user preferences

# From research into high-schools I

Karlsruhe Institute of Technology

$$
R = \begin{array}{c} \\ \text{👤}_1 \\ \text{👤}_2 \\ \text{👤}_3 \\ \text{👤}_4 \\ \text{👤}_5 \end{array}
\begin{array}{c}
\begin{array}{cccccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \text{🎞}_4 & \text{🎞}_5 & \text{🎞}_6 \end{array} \\
\begin{pmatrix}
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & 5 & 4 & 3 & 5 \\
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & \color{red}{?} & 4 & 3 & 5 \\
1 & \color{red}{?} & 4 & 1 & 1 & 2
\end{pmatrix}
\end{array}
$$

$\text{🎞}_1 =$ Batman Returns
$\text{🎞}_2 =$ SpongeBob
$\text{🎞}_3 =$ Simpson
$\text{🎞}_4 =$ Lord of the Rings
$\text{🎞}_5 =$ Miss Congeniality
$\text{🎞}_6 =$ Fast and Furious

$$
M = \begin{array}{c} \\ A \\ C \end{array}
\begin{array}{c}
\begin{array}{cccccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \text{🎞}_4 & \text{🎞}_5 & \text{🎞}_6 \end{array} \\
\begin{pmatrix}
3 & 2 & 1 & 3 & 2 & 3 \\
1 & 2 & 4 & 1 & 1 & 2
\end{pmatrix}
\end{array}
$$

A = Action, C = Comedy

# From research into high-schools I



$$
R = \begin{array}{c} \\ {}_1 \\ {}_2 \\ {}_3 \\ {}_4 \\ {}_5 \end{array}
\begin{array}{c}
\begin{array}{cccccc} {}_1 & {}_2 & {}_3 & {}_4 & {}_5 & {}_6 \end{array} \\
\begin{pmatrix}
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & 5 & 4 & 3 & 5 \\
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & ? & 4 & 3 & 5 \\
1 & ? & 4 & 1 & 1 & 2
\end{pmatrix}
\end{array}
$$

$$
U = \begin{array}{c} \\ {}_1 \\ {}_2 \\ {}_3 \\ {}_4 \\ {}_5 \end{array}
\begin{array}{c}
\begin{array}{cc} \mathrm{A} & \mathrm{C} \end{array} \\
\begin{pmatrix}
1 & 0 \\
1 & 1 \\
1 & 0 \\
1 & 1 \\
0 & 1
\end{pmatrix}
\end{array}
\qquad
M = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{C} \end{array}
\begin{array}{c}
\begin{array}{cccccc} {}_1 & {}_2 & {}_3 & {}_4 & {}_5 & {}_6 \end{array} \\
\begin{pmatrix}
3 & 2 & 1 & 3 & 2 & 3 \\
1 & 2 & 4 & 1 & 1 & 2
\end{pmatrix}
\end{array}
$$

# The mathematical model

$$
R = \begin{array}{c} \\ \text{👤}_1 \\ \text{👤}_2 \\ \text{👤}_3 \\ \text{👤}_4 \\ \text{👤}_5 \end{array}
\begin{array}{c} \begin{array}{cccccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \text{🎞}_4 & \text{🎞}_5 & \text{🎞}_6 \end{array} \\
\begin{pmatrix} 3 & 2 & 1 & 3 & 2 & 3 \\ 4 & 4 & 5 & 4 & 3 & 5 \\ 3 & 2 & 1 & 3 & 2 & 3 \\ 4 & 4 & \mathbf{5} & 4 & 3 & 5 \\ 1 & \mathbf{2} & 4 & 1 & 1 & 2 \end{pmatrix} \end{array}
$$

$$
U = \begin{array}{c} \\ \text{👤}_1 \\ \text{👤}_2 \\ \text{👤}_3 \\ \text{👤}_4 \\ \text{👤}_5 \end{array}
\begin{array}{c} \begin{array}{cc} \mathrm{A} & \mathrm{C} \end{array} \\
\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \end{array}
\qquad
M = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{C} \end{array}
\begin{array}{c} \begin{array}{cccccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \text{🎞}_4 & \text{🎞}_5 & \text{🎞}_6 \end{array} \\
\begin{pmatrix} 3 & \mathbf{2} & \mathbf{1} & 3 & 2 & 3 \\ 1 & \mathbf{2} & \mathbf{4} & 1 & 1 & 2 \end{pmatrix} \end{array}
$$

$$\mathbf{5} = 1 \cdot 1 + 1 \cdot 4$$

$$\mathbf{2} = 2 \cdot 0 + 2 \cdot 1$$

# The mathematical model

$$
R = \begin{array}{c} \\ \text{👤}_1 \\ \text{👤}_2 \\ \text{👤}_3 \\ \text{👤}_4 \\ \text{👤}_5 \end{array}
\begin{array}{c}
\begin{array}{cccccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \text{🎞}_4 & \text{🎞}_5 & \text{🎞}_6 \end{array} \\
\begin{pmatrix}
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & 5 & 4 & 3 & 5 \\
3 & 2 & 1 & 3 & 2 & 3 \\
4 & 4 & \mathbf{5} & 4 & 3 & 5 \\
1 & \mathbf{2} & 4 & 1 & 1 & 2
\end{pmatrix}
\end{array}
$$

$$
U = \begin{array}{c} \\ \text{👤}_1 \\ \text{👤}_2 \\ \text{👤}_3 \\ \text{👤}_4 \\ \text{👤}_5 \end{array}
\begin{array}{c}
\begin{array}{cc} A & C \end{array} \\
\begin{pmatrix}
1 & 0 \\
1 & 1 \\
1 & 0 \\
\mathbf{1} & \mathbf{1} \\
\mathbf{0} & \mathbf{1}
\end{pmatrix}
\end{array}
\qquad
M = \begin{array}{c} \\ A \\ C \end{array}
\begin{array}{c}
\begin{array}{cccccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \text{🎞}_4 & \text{🎞}_5 & \text{🎞}_6 \end{array} \\
\begin{pmatrix}
3 & \mathbf{2} & \mathbf{1} & 3 & 2 & 3 \\
1 & \mathbf{2} & \mathbf{4} & 1 & 1 & 2
\end{pmatrix}
\end{array}
$$

**Matrix factorization** $R \approx U \cdot M$

# Computing a factorization

$$R = \begin{array}{c c} & \begin{array}{cc} \text{🎞}_1 & \text{🎞}_2 \end{array} \\ \begin{array}{c} \text{👤}_1 \\ \text{👤}_2 \end{array} & \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \end{array}$$

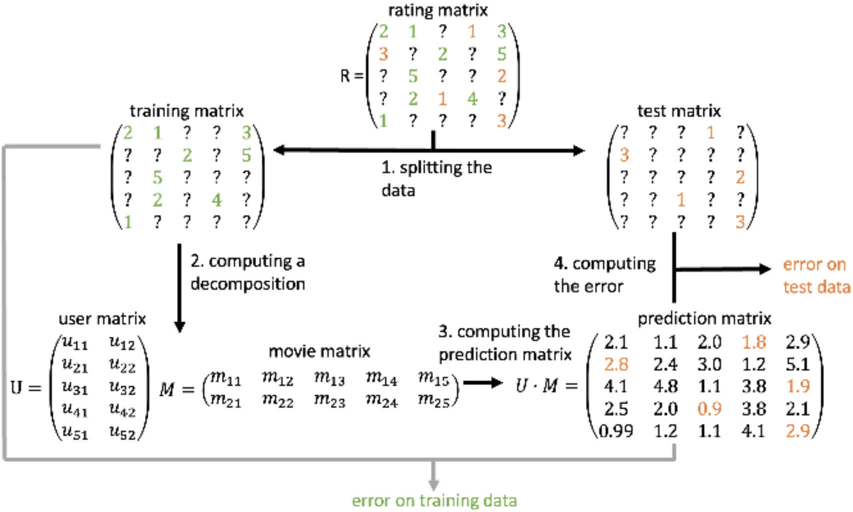
$$U = \begin{array}{c c} & \begin{array}{cc} E1 & E2 \end{array} \\ \begin{array}{c} \text{👤}_1 \\ \text{👤}_2 \end{array} & \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \end{array} \quad \text{and} \quad M = \begin{array}{c c} & \begin{array}{cc} \text{🎞}_1 & \text{🎞}_2 \end{array} \\ \begin{array}{c} E1 \\ E2 \end{array} & \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \end{array}$$

**Find entries of $U$ and $M$, s.t. the error**

$$\Big(r_{11} - (u_{11} \cdot m_{11} + u_{12} \cdot m_{21})\Big)^2 + \Big(r_{12} - (u_{11} \cdot m_{12} + u_{12} \cdot m_{22})\Big)^2 + ...+$$

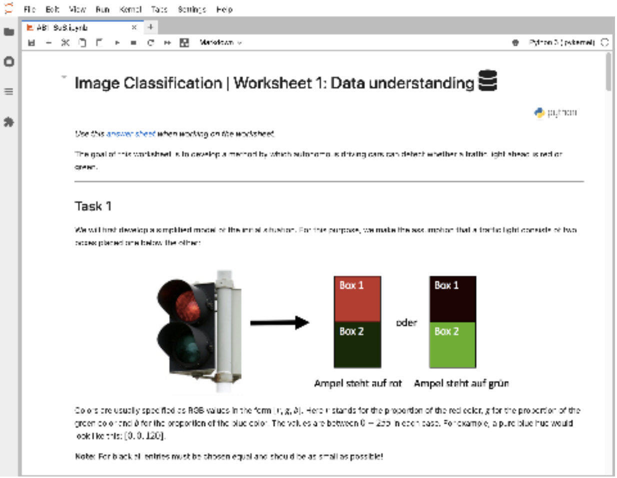
**is minimized** (Optimization as black or white box).

# Validation



rating matrix
$$R = \begin{pmatrix} 2 & 1 & ? & 1 & 3 \\ 3 & ? & 2 & ? & 5 \\ ? & 5 & ? & ? & 2 \\ ? & 2 & 1 & 4 & ? \\ 1 & ? & ? & ? & 3 \end{pmatrix}$$
training matrix
$$\begin{pmatrix} 2 & 1 & ? & ? & 3 \\ ? & ? & 2 & ? & 5 \\ ? & 5 & ? & ? & ? \\ ? & 2 & ? & 4 & ? \\ 1 & ? & ? & ? & ? \end{pmatrix}$$
1. splitting the data
test matrix
$$\begin{pmatrix} ? & ? & ? & 1 & ? \\ 3 & ? & ? & ? & ? \\ ? & ? & ? & ? & 2 \\ ? & ? & 1 & ? & ? \\ ? & ? & ? & ? & 3 \end{pmatrix}$$
2. computing a decomposition
4. computing the error
error on test data
user matrix
$$U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \\ u_{41} & u_{42} \\ u_{51} & u_{52} \end{pmatrix}$$
movie matrix
$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{15} \\ m_{21} & m_{22} & m_{23} & m_{24} & m_{25} \end{pmatrix}$$
3. computing the prediction matrix
prediction matrix
$$U \cdot M = \begin{pmatrix} 2.1 & 1.1 & 2.0 & 1.8 & 2.9 \\ 2.8 & 2.4 & 3.0 & 1.2 & 5.1 \\ 4.1 & 4.8 & 1.1 & 3.8 & 1.9 \\ 2.5 & 2.0 & 0.9 & 3.8 & 2.1 \\ 0.99 & 1.2 & 1.1 & 4.1 & 2.9 \end{pmatrix}$$
error on training data

# Digital tool: Jupyter Notebook




*no programming skills, direct feedback, staggered tips, additional material, no installation of software*

# Netflix Prize & open modeling activities

**Framework:**

- starting point: smaller training and test dataset
- digital tool: non pre-strucutred Jupyter Notebook (CoCalc / Colab)

**Approach of the students:**

- similarities between users via mean rating deviation
- predictions based on $k$ most similar users
- approach is actually in use: **Neighborhood Method**

# From research into high-schools II

$$\begin{array}{ccc} & & \begin{array}{cccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \end{array} \\ \vec{n}_1 = & \text{👤}_1 & \begin{pmatrix} 3 & 1 & 2 & \dots \end{pmatrix} \end{array}$$

$$\begin{array}{ccc} & & \begin{array}{cccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \end{array} \\ \vec{n}_2 = & \text{👤}_2 & \begin{pmatrix} 3 & 1 & ? & \dots \end{pmatrix} \end{array}$$

$$\begin{array}{ccc} & & \begin{array}{cccc} \text{🎞}_1 & \text{🎞}_2 & \text{🎞}_3 & \end{array} \\ \vec{n}_3 = & \text{👤}_3 & \begin{pmatrix} 1 & 4 & 2 & \dots \end{pmatrix} \end{array}$$

**Similarity measures**

- euclidean distance $\|\vec{n}_1 - \vec{n}_2\|$
- cosine similarity $\frac{\vec{n}_1 \cdot \vec{n}_2}{\|\vec{n}_1\| \cdot \|\vec{n}_2\|}$
- Pearson correlation
- ...

**Predicted rating of user *u* for movie *j*:**

(weighted) mean rating of *k* most similar user

(Sawar et al. 2001; Rantzau 2021; Oldenburg 2021; Schönbrodt 2022)

# Experiences with high-school students

- $>$ 200 students from grade 10 (age 15-16) in modeling days / series of lessons
- 15 students in open modeling projects
- diverse ideas, lively discussions, great interest in topic / AI

**What did you learn from participating in the workshop?**

*„That mathematics is more important than I initially assumed.“*

*„That math is included in everyday things like even Netflix.“*

*„I got an interesting insight into how to predict further data from large amounts of data.“*

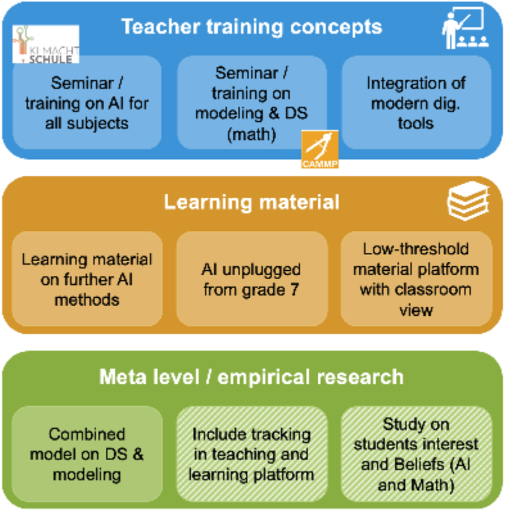
# AI in math high-school education

**Possible!? Reasonable!? Necessary!?**

- **authentic** insight into AI methods is possible
  *connection to school mathematical contents* (*vectors, scalar product, euclidean distance, functions, lines, planes, ...*)
- variety of suitable **real** problems
  *relevance, accessibility, availability of data*
- data-intensive problems allow for **mathematically rich** modeling activities

# Challenges & Outlook

- comprehensive teacher education
- modern digital tools in teacher education
- programming in math education!?



# Thank you!

## Time for discussions.

Access to online learning material?

Suggestions or discussions?

sarah.schoenbrodt@kit.edu

# Literature


- Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
- Blum, W. (1985). Anwendungsorientierter Mathematikunterricht in der didaktischen Diskussion. 32(2),195–232.
- Blum, W. (2010). Modellierungsaufgaben im Mathematikunterricht. Herausforderung für Schüler und Lehrer. Praxis der Mathematik, 34(52), 42–48.
- Chapman, P., Clinton, J., Kerber, R. et al. (2000). CRISP-DM – Step-by-Step Data Mining Guide.
- Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. AI Magazine, 17(3), 37–54.
- Feuerverger, A., He, Y. & Khatri, S. (2012). Statistical Significance of the Netflix Challenge. Stat Sci, 27(2), 202–231.
- Frank, M., Roeckerath, C. & Schönbrodt, S. (im Druck). Einführung. In M. Frank & C. Roeckerath (Hrsg.) Neue Materialien für einen realitätsbezogenen Mathematikunterricht 9. Springer.
- Gerhardt, M., Hattebuhr, M., Schönbrodt, S. & Wohak, K. (im Druck). Aufbau und Einsatzmöglichkeiten des Lehr- und Lernmaterials. In M. Frank & C. Roeckerath (Hrsg.) Neue Materialien für einen realitätsbezogenen Mathematikunterricht 9. Springer.
- Gould, R. (2021). Toward data-scientific thinking. Teaching Statistics, 43, 11–22.

# Literature


- Hofmann, S. & Frank, M. (2022). Teaching data science in school: Digital learning material on predictive text systems. In Proceedings of CERME12.
- Koren, Y., Bell, R. & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. IEEE Comput. Society Press, 42(8), 30–37.
- Niss, M. & Blum, W. (2020). The Learning and Teaching of Mathematical Modelling. Routledge.
- Oldenburg, R. (2021). Big data – small school. In H. Humenberger & B. Schuppar (Hrsg.) Neue Materialien für einen realitätsbezogenen Mathematikunterricht 7. Springer.
- O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality And Threatens Democracy, Crown Books.
- Opel, S., Schlichtig, M., Schulte, C., Biehler, R. et al. (2019). Entwicklung und Reflexion einer Unterrichtssequenz zum Maschinellen Lernen als Aspekt von Data Science in der Sekundarstufe II. In A. Pasternak (Hrsg.), Proceedings zur 18. GI-Fachtagung Informatik und Schule, 285–294.
- Rantzau, L. (2021). Empfehlungssysteme basierend auf Nachbarschaftsmethoden – mathematisch-fachliche Diskussion und Entwicklung digitalen Lernmaterials zur Netflix Challenge für Schüler*innen der Sek II. Bachelorarbeit, KIT
- SAS Institute (2017). Introduction to SEMMA. `https://documentation.sas.com/doc/en/emref/14.3/n061bzurmej4j3n1jnj8bbjjm1a2.htm`

# Literature


- Sarwar, B., Karypis, G., Konstan, J. & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In Proc. 10th Int. Conf. on the WWW, 285–295, ACM.
- Schmidt, L. (2019). Machine Learning: automatische Bilderkennung mit Mathematik?! – Ein Lehr-Lern-Modul im Rahmen eines mathematischen Modellierungstages für Schülerinnen und Schüler der Sek. II. Masterarbeit, RWTH Aachen.
- Schönbrodt, S. (2019). Maschinelle Lernmethoden für Klassifizierungsprobleme – Perspektiven für die mathematische Modellierung mit Schülerinnen und Schülern. Springer.
- Schönbrodt, S. (2022). Optimierungsprobleme in der mathematischen Modellierung: Grundlegende Aspekte und Chancen aus Sicht der Mathematikdidaktik – herausgestellt an aktuellen Problemen aus der Forschung zu künstlicher Intelligenz und erneuerbaren Energien, Dissertation, KIT
- Schönbrodt, S. & Frank, M. (2022). Data Science and Machine Learning in mathematics education: High-school students working on the Netflix Prize. In Proceedings of CERME12.
- Schönbrodt, S., Camminady, T. & Frank, M. (2021). Mathematische Grundlagen der Künstlichen Intelligenz im Schulunterricht. Math Semesterber 69, 73–101.
- Schönbrodt, S., Wohak, K. & Frank, M. (2022). Digital Tools to Enable Collaborative Mathematical Modeling Online. MSEL, 15(1), 151–174.
- Schüller, K., Koch, H., & Rampelt, F. (2021). Data-Literacy Charta. `www.stifterverband.org/sites/default/files/data-literacy-charta_v1_2.pdf`

# Literature


- Sube, M. (2019). Entwicklung und Evaluation von Unterrichtsmaterial zu Data Science und mathematischer Modellierung mit Schülerinnen und Schülern, Dissertation, RWTH Aachen
- Vos, P. (2011). What is authentic in the teaching and learning of mathematical modelling? In G. Kaiser, W. Blum, R. Borromeo Ferri & G. Stillman (Eds.), Trends in Teaching and Learning of Mathematical Modelling (ICTMA14), S. 713–722. Springer.
- Wild, C. & Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. Internat. Stat. Review, 67(3), 223–265.

# AI workshops and teacher trainings

**Topics:** Word predictions, image classification, Netflix Prize, ...
**Focus:** Mathematical modeling, problem-oriented

`www.cammp.online`

**Topics:** AI & medicine, AI & arts, AI & economy, AI & ...
**Focus:** various applications, ethical discussions, history and algorithms of AI / ML

`www.ki-macht-schule.de`