IS4246 Smart Systems and AI Governance Introduction

Professor: Nicholas MacGregor Garcia

The plan for today

Today we will talk about the structure and goals of the course, as well as the broader discussion into which it fits.

Agenda

Course overview

Introductions

Logistics, Grading, and Evaluation

Group Exercise

 Intro to AI, Smart Systems, and Business Applications

Agenda

Course overview

Introductions

Logistics, Grading, and Evaluation

Group Exercise

 Intro to AI, Smart Systems, and Business Applications

Artificial Intelligence needs all of us



https://www.youtube.com/watch?v=LqjP7O9SxOM&ab channel=TEDxTalks

Learning Objectives

- 1. Analyse the use of AI and Smart Systems in different business operations and application fields.
- 2. Develop responsible AI for Smart Systems.
- 3. Evaluate Al and Analytics projects' when addressing data privacy and protection.
- 4. Utilize established Al Governance frameworks for the responsible adoption of Al.
- 5. Identify, mitigate, and manage risks associated with Al-augmented decision-making

What I hope you will learn in this class:



Smart Systems' Role in Business

- Trends
- Economics
- Strategy

Responsible Al and Smart Systems

Design

- Human Centered Design
- Privacy Concerns
- Algorithmic Bias, Fairness and Ethics
- Stakeholder Engagement



Governance Mechanisms

- Emergency response
- Al Safety & Risk management
- Societal Issues

Agenda

Course overview

Introductions

Logistics, Grading, and Evaluation

Group Exercise

 Intro to AI, Smart Systems, and Business Applications

About the Teaching Staff

- Instructor: Nicholas Garcia
 - Office: COM 2 #04-09
 - E-mail: <u>ngarcia@nus.edu.sg</u>
 - Start each subject line with "[IS4246]" in all emails
 - This ensures a prompt response to your inquiries
- Teaching Assistants: Hoang Ton Nu Huong Giang

About Me

My background:

- UC Berkeley (2012)
- Carnegie Melon (2014)
- NYU PhD (2022)

My Research

- Blockchain ecosystem & governance
- Management of Data Science
- Philosophy of Data Science

My Research







Breakout sessions: Introduce yourself



Name



Hobby, interests, or goals



What events, concerns, or opportunities you associate with Al Governance



What skills, knowledge, or opportunities you hope to gain from this course.



Choose one student to summarize and introduce the others to the class.

Breakout sessions: Introduce yourself

One member of each group – introduce the others to the class.

Agenda

Course overview

Introductions

Logistics, Grading, and Evaluation

Group Exercise

 Intro to AI, Smart Systems, and Business Applications

Schedule

Week	Lecture		
1	Introduction to AI / Smart Systems and business applications		
2	Human-Centric Smart Systems		
3	Augmented Analytics		
4	AI-enabled Operations and Stakeholder Management		
5	Privacy Concerns in AI/ML		
6	Algorithmic Bias And Ethics		
	Midterm Break (no exam)		

Schedule

Week	Lecture	
7	AI Governance Structures and Mechanism Design	
8	Risk Management in Smart Systems	
9	AI Safety and Artificial General Intelligence (I)	
10	AI Safety and Existential Risks (II)	
11	Organizational and Societal Implications of Widespread AI	
12	Deep-Dives into Specific Application Areas	
13	Final Project Presentations	

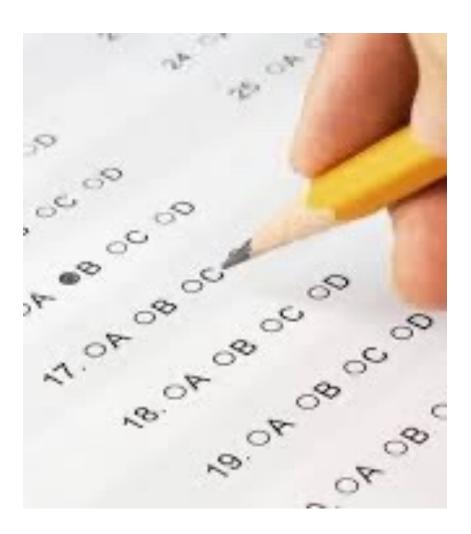
Class evaluation

Component	Weight
Class Participation	10%
Assignments	20%
Group Project and Presentation	30%
Final Exam	40%
Total	100%

Grading

- You will be graded on a curve:
 - ~35% of students can expect to receive A's or A-'s for <u>excellent</u> work
 - 50-60% of students can expect to receive B's for good or very good work
 - 5-15% of students can expect to receive C's or less for <u>adequate or</u> below work
- The actual distribution for this course and your own grade will depend upon how well each of you actually performs in this course.

Participation: Quizzes



- Quizzes will be multiple choice.
- Points for each question will be clearly marked.
- This will help you track your progress during the semester.

Participation: Post-Class Memos & In-Class Discussions

- **Outstanding.** Contribution adds high value to the discussion. The contribution usually involves:
 - 1. Integrating learning across several sessions or topics;
 - 2. Proposing creative yet reasonable ideas; or
 - 3. Offering deep analytical insights or incorporating your own outside knowledge.
- **Good.** Contribution advances the discussion, including responses to questions from the professor. Demonstrates clear grasp of readings. Information presented is relevant to the discussion.
- **Unsatisfactory.** Apparent lack of preparation. Defensive behavior (e.g., aggression or withdrawal).

Assignments: Analysis of Readings and Case Studies



Industry or Government Reports, Case Studies, Articles, Book Chapters, Academic Papers, & Videos



If a reading is marked required, I expect you to have read and understood the articles well, be prepared to talk about it in class and defend a point of view

Assignments: Design of Responsible Al



We will be using iPython notebooks to gain hands on practice and demonstrate your understanding of the concepts being covered in class



If you do not have experience with Python, come see me after class

SMART SYSTEMS AND AI GOVERNANCE

Agenda

Course overview

Introductions

Logistics, Grading, and Evaluation

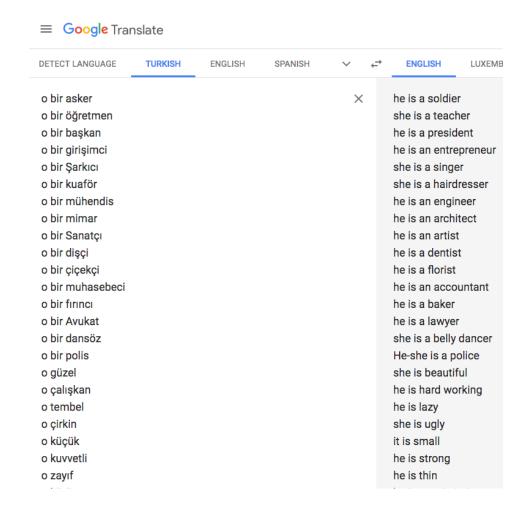
Group Exercise

 Intro to AI, Smart Systems, and Business Applications

What ethical concerns do **you** have around AI?

A few Al Ethics examples...

What's wrong with this picture?



What could the consequences be?



Autonomous machines capable of deadly force are increasingly prevalent in modern warfare, despite numerous ethical concerns. Is there anything we can do to halt the advance of the killer robots?

Let's split into breakout rooms!

10 min: individual reading/slide prep 20+ min: group discussion

Google Slides Link: https://shorturl.at/hqDU0

Students:

On your slide: take 10 minutes to read/watch one link, then add 1 image and 1 discussion question. Be ready to summarize and lead a short discussion!

Agenda

Course overview

Introductions

Logistics, Grading, and Evaluation

Group Exercise

 Intro to AI, Smart Systems, and Business Applications

INTRODUCTION TO AI / SMART SYSTEMS AND BUSINESS APPLICATIONS

Lecture 1 Learning Objectives



Explain the fundamentals of AI and Smart Systems



Describe the various business operations and applications of AI and Smart Systems



Assess the role of Augmented Analytics and Al-driven decision-making

Lecture 1 Learning Objectives



Explain the fundamentals of Al and Smart Systems



Describe the various business operations and applications of AI and Smart Systems



Assess the role of Augmented Analytics and Al-driven decision-making

What is data analytics/AI?

- Hype and buzzwords:
 - Data analytics?
 - Data science?
 - business intelligence?

- Deep learning?
- Big Data?
- AI?

From a <u>business/practical</u> perspective:

- Using IT to extract useful, actionable information from large bodies of data
- Automation

Three goals of data analytics/Al

 Make decisions <u>autonomously</u> or as an <u>aid</u> to a human

- Understand/explore present reality better
- Predict future situations / infer unobservable aspects of the present

Discussion: where is each of these happening? Examples?

What is "AI" (take 1)

- Al is about hunting regularities in data
 - INTERESTING REGULARITIES
 - Strictly speaking, it is not about trying to find patterns known in advance (that would be a task for Pattern Recognition)
 - Most often, it is organized as a procedure of applying analytical tools with the aim of finding something useful or inspiring
 - It usually turns out to be an ITERATIVE PROCESS with a substantial participation of human analysts
 - Improved UNDERSTANDING OF DATA is as important as finding the regularities

Human *is* a factor

- Keeping people in the loop is important, but there are ongoing efforts to reduce the amount of labor required.
- Humans have the capacity to not only make mistakes but can draw insightful conclusions from data, making remarkable breakthroughs.

Human *is* a factor

- Augmented Analytics tools streamline the data analysis process by automatically predicting which data transformations, volume, analysis, and visualizations would be most effective.
- This makes it easier for non-technical users to leverage the organizations' data with data, allowing them to explore more advanced data topics.

What is "AI" (take 2)

- Finding relationships in data using smart statistics and computer science
 - SMART COMPUTER SCIENCE allows massive data sets to be analyzed
 - SMART STATISTICS accurately identifies relationships
 - SMART ALGORITHMS utilize results to improve decision making

- Academics because of the interesting database, statistics, machine learning, and algorithmic challenges.
 - Conferences on Knowledge Discovery and Data Mining
 - International Conferences on Machine Learning (ICML)
 - ... many many others ...

- Vendors because of a huge market potential
 - Database companies (e.g. Oracle)
 - Hardware companies (e.g. IBM)
 - Consulting companies (e.g. Deloitte)
 - Software companies (e.g. SAS, numerous machine learning startups)

Businesses

 including organizations that have been collecting data for years because of their operating policy, but who have not made a substantial use of it

Businesses

 including organizations that have been collecting data for years because of their operating policy, but who have not made substantial use of it

> A quote from Our Friendly Customer: Yeah, we have plenty of DATA. But what we really need is INFORMATION...

AI / Data mining key components

Availability of data

Computational power

"Instrumentation" of companies

Statistics

Need for information/knowledge

Maturity of Artificial Intelligence and Machine Learning

Need for Knowledge: Marketing questions one can answer with Data Mining

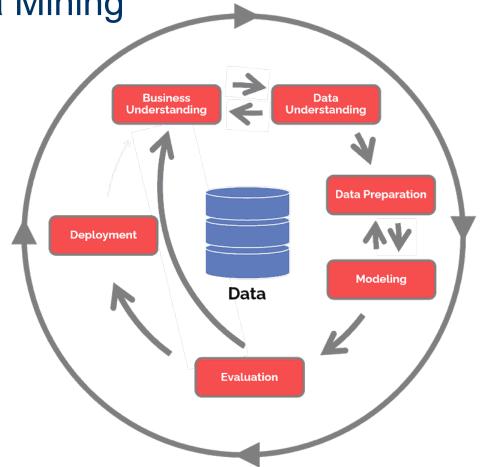
- Which new products should I recommend or advertise to a person?
- How should I segment my products?
- How should I segment my customers?
- Which cross-promotion strategy is likely to be the most profitable?

Data Mining cycle (the analyst's perspective)

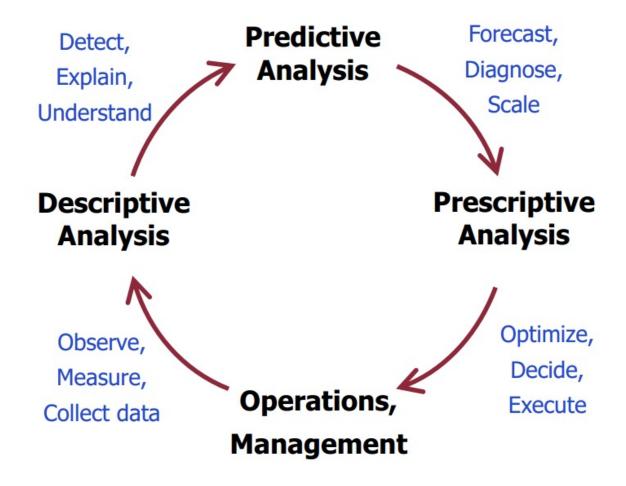
- (1) You are presented with a set of (sometimes mysterious) data (if your luck is average, it is cryptic or, perhaps, misleadingly labeled)
- (2) Sometimes you are roughly informed what the customers are up to, or what do they want from data (otherwise you are asked to 'find something useful' in it)
- (3) According to (2) and (1) you choose and apply some analytical method(s) to the data
- (4) After observing and evaluating the outcomes of step (3) you may:
- Re-define your initial hypotheses, change your views, change/refine your analytical methods and/or their settings and repeat from (3), or
- Talk to the customers (get back to (2)) in order to gather more of background knowledge. This might help you better understand their objectives and choose a more adequate analytical tool, plus – after giving their data a try – you are likely to find some common platform of discussion with them if that has not developed yet;
- Focus your attention on some particularly interesting aspects of the data and the analyzes conducted so far, adjusting the methodology, or
- Conclude the loop and report the results to the customer!

Data Mining cycle: Best practice standards

 CRISP-DM: CRoss Industry Standard Process for Data Mining



The Cycle of Analytics



How does the business get done?

A. Development of own methodologies and/or proprietary implementations

B. Using software tools available on the Market

C. Hiring professionals to do the job for you

Data and corporate strategy

- "We have come out on top in the casino wars by mining our customer data deeply, running marketing experiments and using the results to develop and implement finely tuned marketing and services strategies that keep our customers coming back."
 Gary Loveman, CEO, Harrahs
- "For every leader in the company, not just for me, there are decisions that can be made by analysis. These are the best kind of decisions. They're fact based decisions."
 Jeff Bezos, CEO, Amazon
- "It's all about collecting information on 200 million people you'd never meet, and on the basis of that information, making a series of very critical long-term decisions about lending them money and hoping they would pay you back." Rich Fairbank, founder and CEO of Capital One

The tech behind the hype: the basics

Key tasks that lie behind analytics/Al:

Optimize	Choose the that has the highest (or lowest)
Classify	Into which category of does this belong?
Predict	Which of these scenarios is most likely?
Generative Al	Create a new based on the input data
Pattern detection	Show me which often co-occur Group together based on their similarity
Visualization	Display massive number of in a way a human can more easily analyze

Model learning scenarios

Supervised learning

(alternative names: by example, by being taught, by imitation) typical for regression and classification-type tasks, when training data is readily available for empirical evaluation and adaptation of the models

Unsupervised learning

(self-organization, no specific output dimension(s)) typical for clustering and density estimation

Semi-supervised learning

(only some training data has known output values)

Reinforcement learning

(by experimentation, trial-and-error, with delayed rewards)

typical for control and optimization problems, also in the context of
classification and regression when training data is not readily available or
comes in short supply

Supervised and unsupervised methods

Supervised methods:

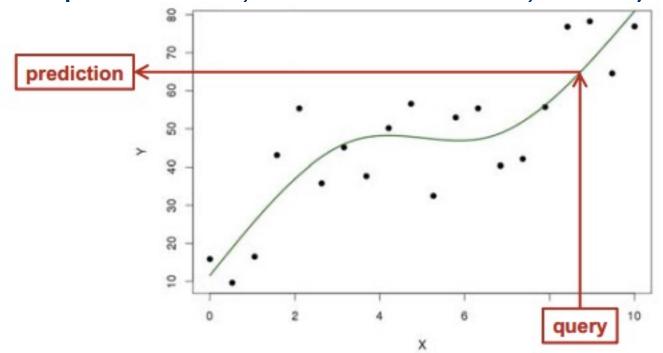
- aim to do inference / prediction
- can check the prediction strength by seeing if it can pick the "right" answers
- model is **built** using historical "training data" for which the answer is known
- model is deployed on new data (for which answer is not yet known)

Uses:

- Automation
- Recommendations

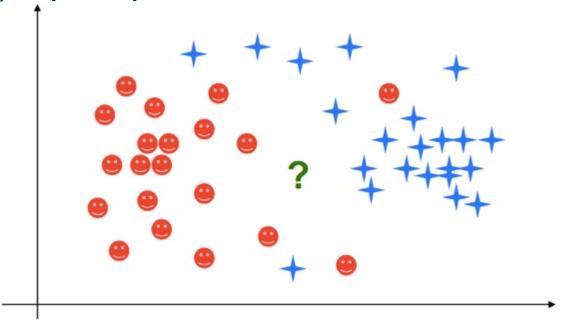
Supervised methods: Regression

- Input(s) are numeric, output(s) are numeric
- Plenty of approaches (simple linear, nonlinear, non-parametric, neural networks, etc...)



Supervised methods: Classification

- Output is symbolic, inputs are numeric or symbolic or mixed
- An example of a binary classification problem in a (2-D) input space is shown:



Supervised and unsupervised methods

Unsupervised methods:

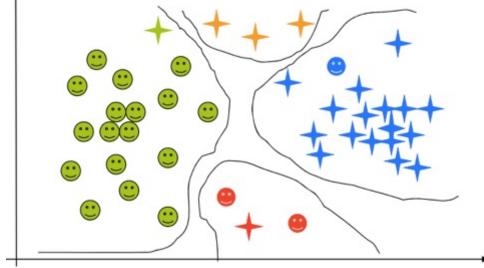
- not trying to get an "answer" to a question
- aim to help human analyst or other system better characterize/understand dataset

Uses:

- "Learning" relationships between things
- Compresses information into "higher-level" features (improves performance on supervised tasks)

Unsupervised method: Clustering

- Input(s) are numeric or symbolic or mixed
- No outputs (labels are unknown)
- How many are there distinctive groups of data and what are the shapes and locations of the areas they populate?
- Sort of



Unsupervised method: Clustering Example

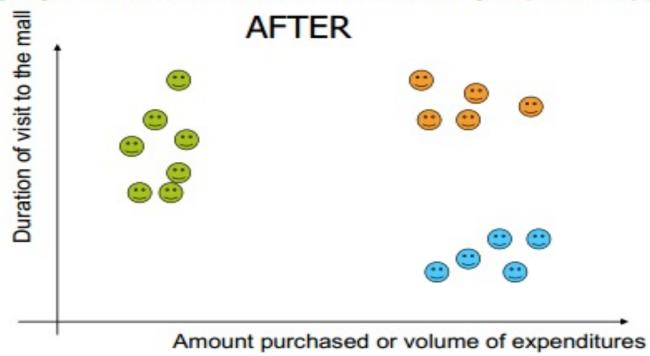


Amount purchased or volume of expenditures

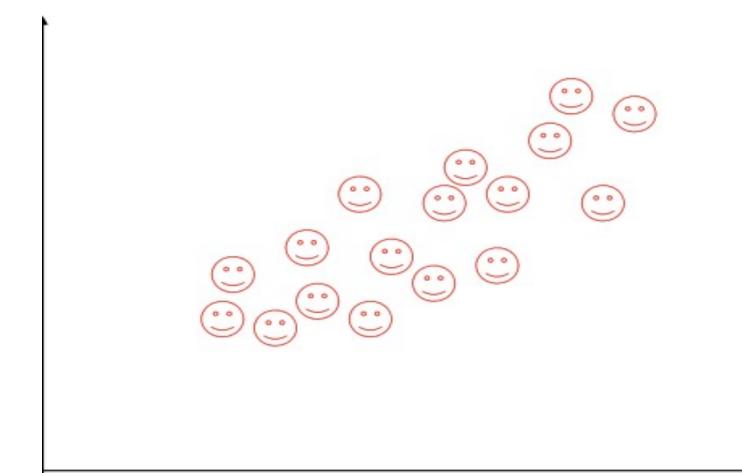
Unsupervised method: Clustering Example

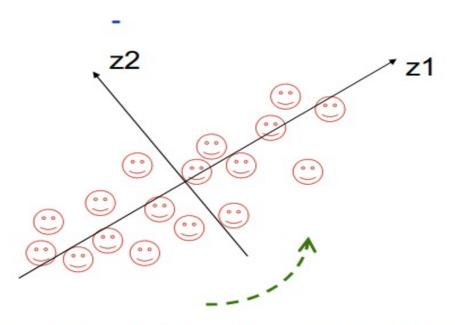
Grouping shoppers

clustering algorithm has discovered three distinct groups of shoppers



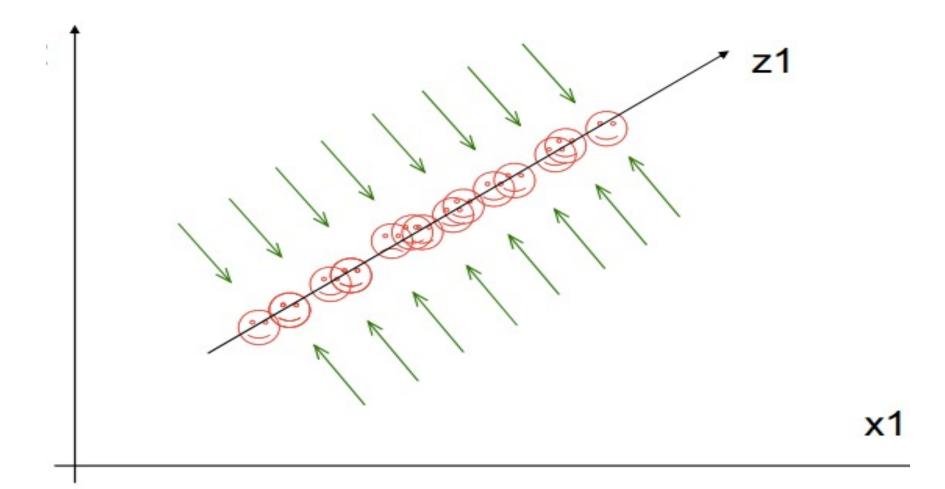
Unsupervised Method: Principal Component Analysis





Then, we rotate the system so that one of the axes aligns with the direction of the greatest variance

- After transformation, we could skip z2 and keep only z1
- By skipping z2 we surely lose some information, but at the smallest price (in the linear reconstruction error sense)



Principal Component Analysis

While clustering can find groups of "entities", PCA can find which "attributes" group together. These grouped attributes can be "more intuitive" and general than the fine-grained data.

Principal Component Analysis Example

PCA from combined Yelp and US Census data for every zipcode in the USA.

PC1: Wealth level of the zip:

- Above poverty line
- Income levels
- Occupation: management, business, science, or art
- US Citizenship
- Speak only English

PC2: Rural Economy:

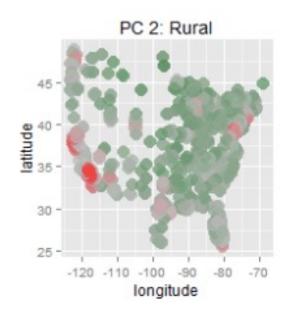
- Fewer businesses: Home services, auto repair, airports, colleges, taxis, etc.
- Short commute
- Lower Income
- Less foreign born

PC3: Car Culture (suburb):

- More vehicles available
- Less renters
- Less public transit
- Don't commute during rush hour
- Less walking, public transit, Bicycle

PC4: Blue collar:

- Commute in a car, truck, or van
- Not White
- Male
- Low Income
- Service Workers
- Mainly drive alone to work



Principal Component Analysis Example

PCA from combined Yelp and US Census data for every zipcode in the USA.

PC5: Young Suburbs / Immigrants / urban periphery:

- Leave for work early
- Long Commute
- Fewer bars, restaurants, colleges, hotels
- Recently moved there
- Fewer elderly, more youth
- Jobs in construction, transportation, and warehousing
- More foreign born & spanish speakers

PC6: Car-Based Big Cities: (The cities with the highest score are Atlanta, LA, Miami, etc.)

- Long commute
- Moderately low income (\$35k-\$50k)
- More bars, schools, colleges
- Less agriculture

PC7: Unsatisfied Shoppers and Yuppies

- Low rated businesses
- More public transit
- Short commute
- Less females
- More young people

PC8: Government town:

 Jobs with Federal government, military

Fundamental modeling tasks: Summary

Classification

Inputs → Class label

Regression

Inputs \rightarrow Number

Clustering

All attributes → Categories

Density estimation

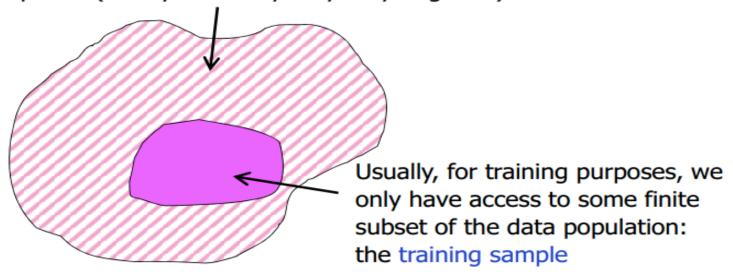
All attributes → Model of data distribution

Structural learning

All attributes → Structure of relationships in data

Generalization

During its life, our system may be exposed to a whole lot of data points (it may be a very very very large set)



But, we want our system to perform well over the whole domain, including the data never actually seen during training

We want it to **generalize** well

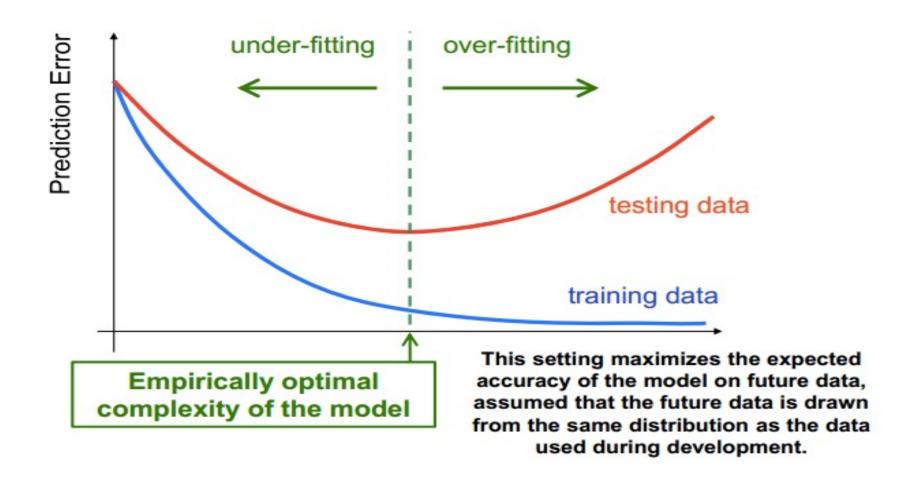
How to arrange for that?

Generalization

 So, for practical reasons, we are primarily interested in using our trained models to predict outcomes for new, yet unseen input patterns

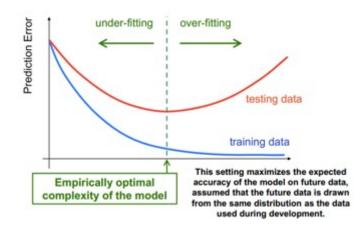
- How do you measure the chance that the system produces good results on unseen data?
 - "Hold out" some data from the training and see if the model correctly predicts it

Generalization curve



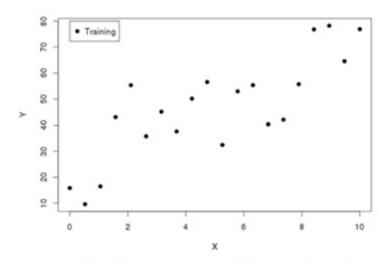
Model Complexity

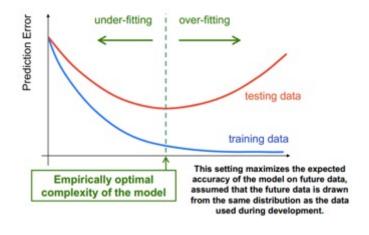
- Higher complexity:
 - More variables
 - Not "forced" to be simple



Model Complexity

- Higher complexity:
 - More variables
 - Not "forced" to be simple

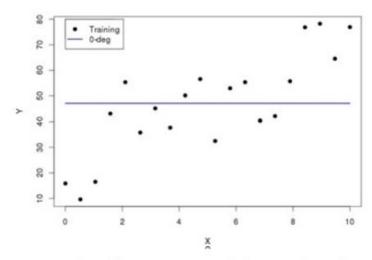


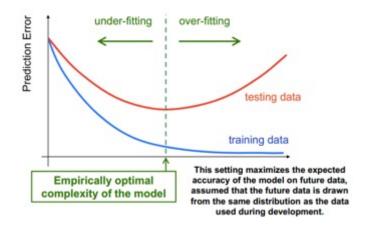


We are supposed to fit some model to noisy data

 What would be the simplest kind of function (model type) we could try on this data?

- Higher complexity:
 - More variables
 - Not "forced" to be simple

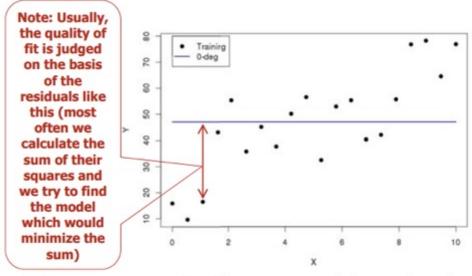


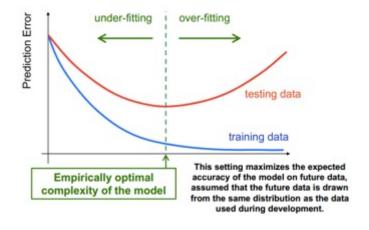


We are supposed to fit some model to noisy data

 What would be the simplest kind of function (model type) we could try on this data?

- Higher complexity:
 - More variables
 - Not "forced" to be simple

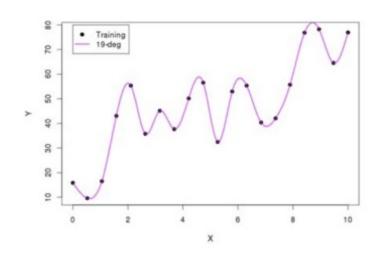


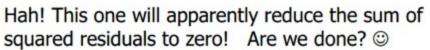


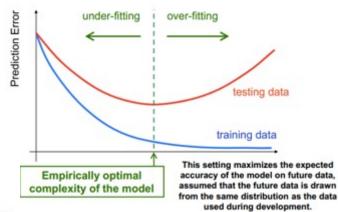
We are supposed to fit some model to noisy data

 What would be the simplest kind of function (model type) we could try on this data?

- Higher complexity:
 - More variables
 - Not "forced" to be simple

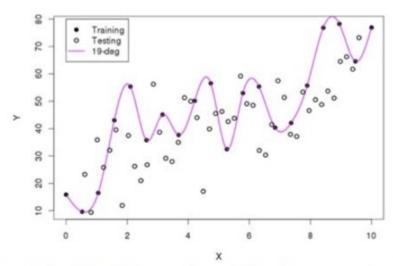


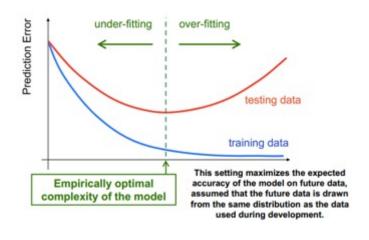




RMSE: f₁₉: 0.0

- Higher complexity:
 - More variables
 - Not "forced" to be simple

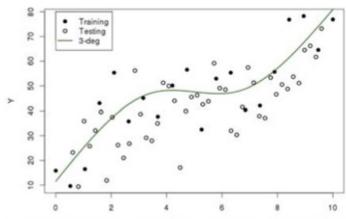


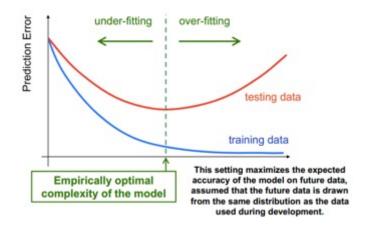


Unfortunately, the things do not look so great when we plot the subset of data that we set aside for testing...

What has just happened?

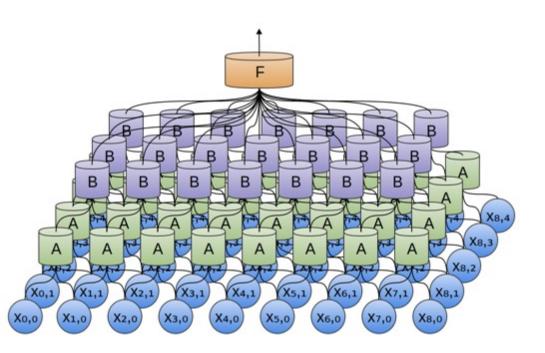
- Higher complexity:
 - More variables
 - Not "forced" to be simple

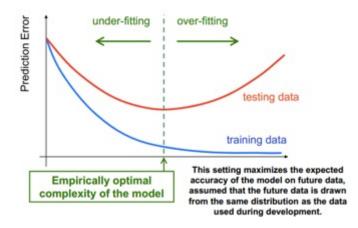




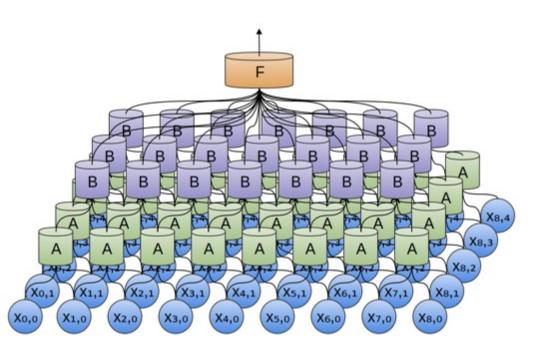
- By tracing the test-set error, we were able to find the model that appears optimal for mitigating future risk for this data
- There may be even more reliable models out there, but we only tried one family (polynomials)

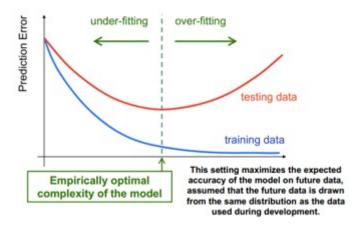
- Higher complexity:
 - More variables
 - Not "forced" to be simple

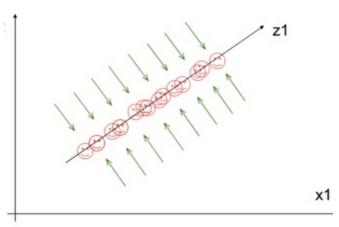




- Higher complexity:
 - More variables
 - Not "forced" to be simple

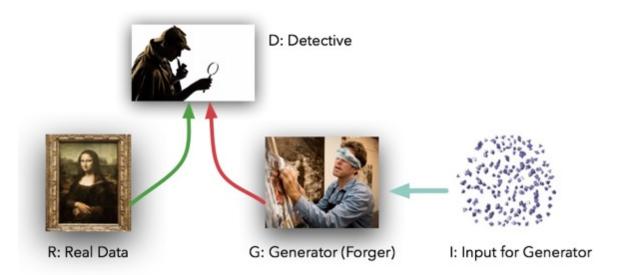






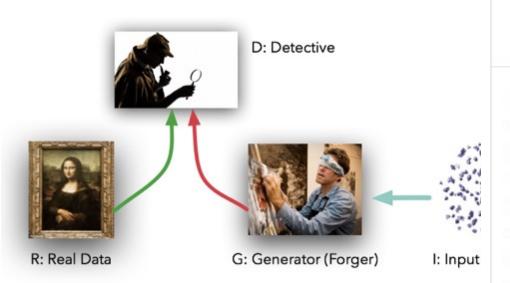
Neural Networks are getting better at Generating Data

 Techniques such as Generative Adversarial Neural Networks (GANs)



Neural Networks are getting better at Generating Data

 Techniques such as Generative Adversarial Neural Networks (GANs)





Deepfake



Deepfake is a technique for human image synthesis based on artificial intelligence. It is used to combine and superimpose existing images and videos onto source images or videos using a machine learning technique known as generative adversarial network. The phrase "deepfake" was coined in 2017. Wikipedia

Neural Networks are getting better at Generating Data FaceApp Mind-blowing effects



Make them smile



Look younger



Meet your future self



Change your style



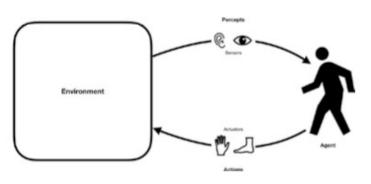


Figure 2-1. An agent acts with its environment through sensors and actuators. Humans are agents that perceive through sensors like eyes, ears, nose, tongue, or touch. We act through actuators like hands and legs

Smart systems

- Incorporatesensing,actuation, andcontrol functions
- Make decisions
 based on available
 data in a
 predictive or
 adaptive manner

Table 2-1. Examples of agents and environments with their sensors, actuators, and goals

Goal	Agent	Environment	Sensors	Actuators
Clean a room	Robotic vacuum cleaner	Room, dust	Camera, infrared	Vacuum suction, brush
Find something on the Internet	Search engine	Websites on the Internet	Keyboard input	Display of search results
Sort vegetables	Vegetable sorting machine	Pile of vegetables	Camera	Robotic arms
Email spam detection	Spam detection algorithm	Emails in the inbox	Incoming email, other user information	Mark as span
Convert speech to text	Speech-to- text engine	Audio file	Microphone	Display of tex transcription
Predict the price of a house	Price estimation algorithm	Specifications of houses like location, size, number of rooms, price of similar houses, etc.	Keyboard input or mouse click on a listing	Display of estimated price

Smart systems

- Incorporatesensing,actuation, andcontrol functions
- Make decisions
 based on available
 data in a
 predictive or
 adaptive manner

Types of Control

- Closed-Loop (aka Feedback)
 - The action from the smart system is dependent on the process' output
 - Can eventually correct any errors that it makes
- E.g., Cruise Control on a Car

Reinforcement learning

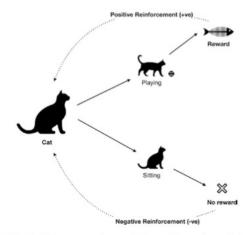
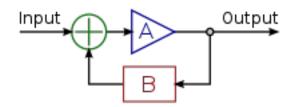
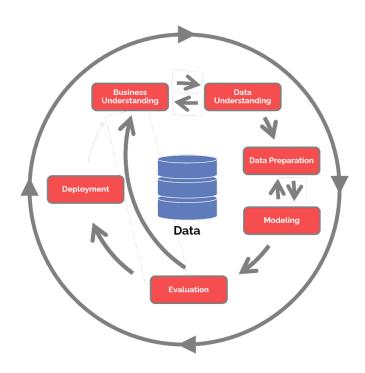


Figure 2-4. Training a cat using positive and negative reinforcement



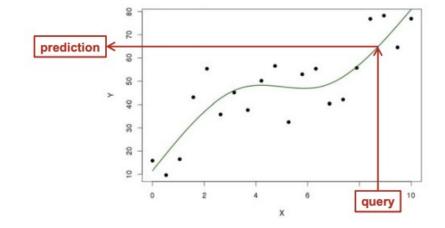
Types of Control

- Closed-Loop (aka Feedback)
 - The action from the smart system is dependent on the process' output
 - Can eventually correct any errors that it makes
- E.g., Cruise Control on a Car



Types of Control

- Open-Loop (aka Feed Forward)
 - The result is known to be approximately sufficient under normal conditions without the need for constant feedback
 - Cannot correct any errors that it makes or correct for outside disturbances



Lecture 1 Learning Objectives



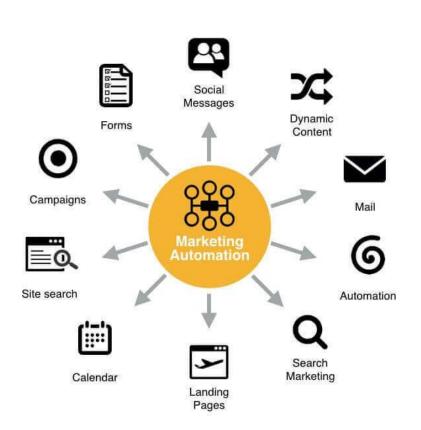
Explain the fundamentals of AI and Smart Systems



Describe the various business operations and applications of Al and Smart Systems



Assess the role of Augmented Analytics and Al-driven decision-making



Marketing Automation:

- Combines multiple marketing tools (CRM, email marketing, social media marketing)
- Target customers with automated messages across email, web, social, and text.
- Personalizes customer experiences & predicts customer trends & behavior

Customer Service Automation:

- Automates call routing & customer segmentation
- Provides automated customer support
- Identifies customer preferences





Logistics:

- Optimizes routing & reduces transport time
- Monitors real-time traffic conditions & detects potential issues
- Improves delivery efficiency

Financial Services:

- Provides financial analytics& monitoring
- Detects fraud
- Automates loan approval processes
- Used in investment & portfolio management



Manufacturing:

- Automates factory tasks & processes
- Monitors manufacturing quality
- Detects potential defects
- Identifies environmental anomalies
- Optimizes inventory management



Lecture 1 Learning Objectives



Explain the fundamentals of AI and Smart Systems



Describe the various business operations and applications of AI and Smart Systems



Assess the role of Augmented Analytics and Al-driven decision-making

Role of Smart Systems in Business

- Intelligent Automation is the combination of AI and automation to solve complex problems quickly and accurately.
- It offers cost savings, increased performance, improved customer experience, and higher service quality.
- It can be applied to a range of business functions including marketing, finance, customer engagement, and accounting.

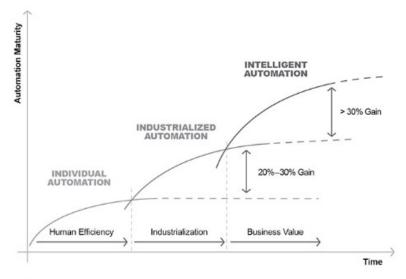


FIGURE 1.1 Conventional to Maturity Automation Journey

Levels of Automation Maturity



Al Driven

- The Five levels are distinguished by managerial mindset.
- Achieving value depends on growing data, skills & foundations.



Data Driven



Robotic Process Automation (RPA)
Driven



Process Driven



Tools-Driven

Tools Driven

- Automation efforts in the first level of maturity are often fragmented and independent from each other
- Benefits yield from individual solutions, but value is limited due to these efforts being in 'pockets'
- Early successes encourage further automation efforts
- Skills and mindset developed at this level set the foundation for further automated capabilities.



Process Driven

- Second-level automation maturity: streamlining or automating whole process across teams or functions, eliminating unnecessary steps
- Process improvement methods: e.g., Lean & Six Sigma
- Banking examples: remote banking for deposits, transfers, account balance, credit card details



Robotic Process Automation (RPA) Driven



RPA automates rules-based processes involving structured data for tedious, painful manual tasks.

- </>>
- Integrates with various other applications.

e<

Popular applications include employee onboarding, sales streamlining, and customer interaction.

 \searrow

Can integrate AI, for example sorting incoming emails by by reading them

Companies reaching this level of automation tend to set aside more resources toward automation and leveraging data

Data Driven - Smart Insights



Data-driven automation can make processes more agile and predictive



Data-driven automation can generate insights to fuel intelligent technologies



Data needs to be of high-quality, transparent, and trustworthy to be mobilized effectively

Intelligence Driven – Smart Automation



Al used to drive revenue growth by augmenting human workers



Al for decision-making, offering unique customer experiences, and optimizing performance



Model-driven AI can improve customer interactions and learn from mistakes



Example: Insurers using Al for claim reimbursements, providing more tailored and cheaper policies, and detecting fraud

Key Questions we'll explore in this class

- How can AI and robotics enhance productivity in various industries?
- What are the potential impacts of AI and robotics on labour markets and employment?
- How do algorithmic decision-making and bias impact the use of AI and robotics?
- What are the organizational design implications of integrating AI and robotics into workflows?

Thank You!