# What is Machine Learning? INFO-4604, Applied Machine Learning

University of Colorado Boulder

August 29-31, 2017

Prof. Michael Paul

# Definition

Murphy:

 "a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data"

# Definition

Murphy:

- "a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data"
- predict = guess the value(s) of unknown variable(s)
  - (not necessarily prediction of future... c.f. *forecasting*)
- future data = data you haven't seen before

# **Types of Learning**

- Supervised learning
  - Goal: Prediction
- Unsupervised learning
  - Goal: Discovery
- Reinforcement learning

# Supervised Learning

Learn how to *predict* an output from a given input.

- Given a photo, identify who is in it
- Given an audio clip, identify the song
- Given a patient's medical history, estimate how likely they will need follow-up care within a month

# Supervised Learning

Two types of prediction:

- Classification
  - Discrete outputs (typically categorical)
- Regression
  - Continuous outputs (usually)

If you need to brush up on these definitions, read Ch. 1 of *OpenIntro Statistics*.

# Classification

#### Document classification

- Is this email spam?
- Is this tweet positive toward this product?
- Is this review/article real?
- Image classification
  - Is this a photo of a cat?
  - Which letter or number is written here?
- Object recognition
  - Identify the faces in this image
  - Identify pedestrians in this video

# Classification

A classification algorithm is called a **classifier** 

Classifiers require examples of inputs paired with outputs

Called training data

Classifiers learn from training examples to map input to output

• Then when a classifier encounters new data where the output is unknown, it can make a prediction



(attributes, measurements, dimensions)





Α	В	С	Prediction
13	Ν	Ν	Υ
15	Ν	Υ	Ν
16	Ν	Ν	Υ
22	Ν	Υ	Ν
28	Υ	Ν	Υ
41	Ν	Ν	N

Α	В	С	Prediction
14	Ν	Υ	?
15	Ν	Ν	?
17	Υ	Υ	?
26	Ν	Υ	?
30	Υ	Ν	?
30	Ν	Ν	?

Α	В	С	Prediction
14	Ν	Υ	Ν
15	Ν	Ν	Υ
17	Υ	Υ	Υ
26	Ν	Υ	Ν
30	Υ	Ν	Υ
30	Ν	Ν	N

What are we predicting? "Will this consumer like the new Taylor Swift single?"

- What are the features? A = age of consumer (years)
- B = did this person purchase Taylor Swift's previous album? (yes/no)
- C = does this person like Kanye West? (yes/no)

Age	Previous Purchase	Likes Kanye	Likes New TSwift
13	Ν	Ν	Υ
15	Ν	Υ	Ν
16	Ν	Ν	Υ
22	Ν	Υ	Ν
28	Υ	Ν	Υ
41	Ν	Ν	Ν

# Let's build a classifier: takeaway

Lots of rules match the original data

- Most rules won't work on new data
- Need to be able to generalize

This is hard to do without knowing what the variables mean

- A machine learning algorithm won't know what they mean, either (unless you tell it)
- Some heuristics: use rules with lots of evidence; use rules that are simple

# Supervised Learning

Recipe for supervised machine learning:

Pattern matching + generalization

# Supervised Learning

Two types of prediction:

- Classification
  - Discrete outputs (typically categorical)

#### Regression

Continuous outputs (usually)

# Regression



Linear regression with one input variable

# Regression

#### Examples:

- Predicting how much money a movie will make
- Forecasting tomorrow's high temperature
- Estimate someone's age based on their face
- Rate how strongly someone likes a product (e.g., in a tweet)

# **Types of Learning**

- Supervised learning
  - Goal: Prediction
- Unsupervised learning
  - Goal: Discovery
- Reinforcement learning

# **Unsupervised Learning**

Finding "interesting" patterns in data

- Not trying to predict any particular variable
- No training data
- Maybe you don't even know what you're looking for

#### Example: anomaly detection

• Trying to identify something unusual (e.g., fraud) but you don't know what it looks like

# **Unsupervised Learning**

**Clustering** is an unsupervised learning task that involves grouping data instances into categories

 Similar to classification, but you don't know what the classes are ahead of time



# **Unsupervised Learning**

Example: movie recommendation (the Netflix problem)

 Clustering can be used to put people into different groups based on the kinds of movies they like.

**Interest Group 3:** 

Trainspotting Fargo Pulp Fiction Clerks Interest Group 18: Mary Poppins Cinderella The Sound of Music Dumbo

Interest Group 8: Pretty Woman Mrs. Doubtfire Ghost Sleepless in Seattle

From Hoffman (2004) "Latent Semantic Models for Collaborative Filtering."

#### Classification



#### Regression



Clustering



# Semi-supervised Learning

Combines both types of learning

Really just a special case of supervised learning

 You have a specific prediction task, but some of your data has unknown outputs

# **Types of Learning**

- Supervised learning
  - Goal: Prediction
- Unsupervised learning
  - Goal: Discovery
- Reinforcement learning

# **Reinforcement Learning**

Setting:

- an agent interacts with an environment
- actions by the agent lead to different states of the environment
- some states will provide **rewards**

Learning goal is to maximize rewards. Used to learn models of how to behave, more complex than just input $\rightarrow$ output

# **Reinforcement Learning**

Most commonly used for creating robots and automated vehicles

Can also learn to play games



Some uses in more traditional machine learning tasks by creatively defining what the agent and environment are

#### Pause



(attributes, measurements, dimensions)

Each data point (i.e., each "thing" you are classifying/regressing/clustering) is called an **instance** 

- Alternative name: **observation**
- Also called examples or samples when used as training data in supervised learning

In a data set, each row corresponds to an instance.

The "input" variables are called **features** 

- Alternative names: attributes, covariates
- Also referred to as the **independent** variables

In a data set, each column corresponds to a feature. (Except for the last column, which is the output.)

The list of feature values for an instance is called the instance's **feature vector** 

The value of the "output" variable (the "thing" you are trying to predict) is the **label** 

• Also called the **dependent** variable

In a data set, this is the final column. (Unless there is more than one label, which is a setting we will consider later in the course.)

In classification, the possible values the labels can have are called **classes** 



(attributes, measurements, dimensions)

In supervised learning:

- a training instance is a feature vector paired with a label
- the **training data** (sometimes **labeled data**) is the table of all training instances

In unsupervised learning, the data set contains feature vectors but no labels (sometimes called **unlabeled data**)

A **prediction function** is what you get at the end of learning

- Sometimes called a predictor (but features are also sometimes called predictor variables, so this can get confusing)
- Sometimes called a hypothesis

A **classifier** is what you call a prediction function if you are doing classification.

#### Example of a simple prediction function: y = .17x + 5



Where does this function come from? Need to learn it so that it is accurate. What is accurate?

Need to define the **error** or **loss** of a prediction function.

- For classification, this is usually the probability that the classifier will output the correct label.
- For regression, this is usually measured by how far away the predicted value will be.

There is some hypothetical measure of how well a classifier will do on all data it might encounter (the **true error** or **risk**)

But there's probably no way to measure that... usually you can only measure the error or loss on the training data, called the **training error** 

Alternatively: empirical error/risk

Goal of machine learning is to learn a prediction function that minimizes the (true) error.

Since true error is unknown, instead minimize the training error.



From: https://xkcd.com/1122/

. . .

. . .

. . .

with stati	BLEM Ements like	NO ONE HAS BEEN Elected President	NO INCUMBENT HAS EVER BEEN REELECTED.	NO ONE WITHOUT FALSE TEETH HAS	No challenger has Beaten an incumbent.
"NO <party> ( WON THE ELECTION</party>	"ANDIDATE HAS WITHOUT <strate>" X8</strate>	BERORE.	\$ <sup>′</sup>	BECOME PRESIDENT:	۶ź
"NO PRESID REELECTED UNDE	∽ ENT HAG BEEN ER ≪ORCUMSTANCES> <sup>™</sup>	λ	1\	λ	Ά
		BUT WASHINGTON WAS.	UNTIL WASHINGTON.	BUT ADAMS DID.	BUT JEFFERSON DID.
1904	[1808]	1812	1816	1820)	1824
No incombent has Beaten a Challenger.	NO CONGRESSMAN HAS EVER BECOME PRESIDENT	NO ONE CAN WIN WITHOUT NEW YORK.	NO CANDIDATE WHO DOESN'T WEAR A WIG	NO ONE WHO WEARS PAINTS INSTEAD OF REFERENCE (ON INF	NO ONE HAS EVER WON WITHOUT A
₽́	Q (	Q,	Q	REELECTED.	Q
Ά	ĺΛ	1 λ	Ά	/\	λ
UNTIL JEFFERSON.	UNTIL MADISON.	BUT MADISON DID.	UNTIL MONROE WAS.	BUT MONROE WAS.	J.Q. ADAMS DID.
[1829]	1832	[1836]	1940	[1844]	1949
UNLY PEOPLE FROM MASSACHUSETTS AND VIRGINIA CAN WIN	WHO GET REELECTED	NEW YORKERS AWAYS LOSE.	NO ONE OVER 65 HAS WON THE PRESIDENCY.	NOONE WHO'S LOST HIS HOME STRITE HAS WON.	AS GOES MISSISSIPPI, SO GOES THE NATION.
			0		
۲ ۲	۲X	l K	Γ I	Γ I	K I
		IN THE REPORT			
UNIIL JHUSON DID.	UNIIL JHUKSON.	1860	UNUL MAKKOW UD.	OUL POLK OID.	UNIIC 1010.
NEW ENGLAND	NO ONE CAN DECOME	NO ONE OVER 6'5"	NO ONE WITH	NO ONE CAN BE	NO ONE WITH A
DEMOORATS CAN'T WIN.	PRESIDENT WITHOUT GETTING MARRIED.	CAN GET ELECTED.	a Beard Has Been Reelected.	PRESIDENT IF THEIR PARENTS ARE AUVE.	BEARD HAS BEEN REELECTED IN PEACETIME.
~ <b>?</b>	Ŷ	Ŕ	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	Ŕ	8
Λ	Λ	Δ	λ	Λ.	Λ
UNTIL PIERCE DID.	UNITIL BUCHANAN DID.	UNTIL LINCOLN.	BUT LINCOLN WAS.	UNTIL GRANT.	UNTIL GRANT WAS.
NOONE CAN WIN	AS COES (AURINANIA	CANDIDATES NAMED	NO SITTING PRESIDENT	NO FORMER PRESIDENT	TALL MIDWESTERNERS
A MAJORITY OF THE POPULAR VOTE AND	SO GOES THE NATION.	"JAMES" CAN'T LOSE.	HAS BEEN BEATEN SINCE THE CIVIL WAR.	HAG BEEN ELECTED.	ARE UNBEATABLE.
STILL LOSE.	2	Ŕ	Qź	۶ź	۶ź
ή	Ά	Ά	1/	Ά	Λ
TILDEN DID.	UNTIL IT WENT HANCOCK.	UNTIL JAMES BLAINE.	CLEVELAND WAS.	UNTIL (LEVELAND.	BRYAN WASN'T.
NO REPUBLICAN	NO ONE UNDER US	NO PERINGULANI LINA	APTER UNCON REAT	NO DEMOCRAT HAS	1920
SHORTER THAN 5'8"	HAS BEEN ELECTED.	HASN'T SERVED IN THE	THE DEMOCRATS WHILE	WON WHILE LOSING WEST VIRGINIA.	NO INCUMBENT SENATOR HAS WON.
THIS OLDN NEELECTER.		MILLIAKY HAS WON.	STORE OF TOUR OF WHITE		
2	Ŕ		No mustache, the only Democrats who can	8	<u>م</u>
	Â		No MUSTACHE, THE ONLY DEMOCRATS WHO CAN UN HAVE A MUSTACHE. WITH NO BEARD.	Ň	Ŕ
UNTIC MCKINLEY WAS	ROOSEVIELT WAS.	MILINARY HAS WON.	No mustrate, the only Democrate who can un have a mustrate With No Beard. Wilson had nether.	Wilson DD.	UNTIC HARDING.
UNTIL MCKINEY WAS.	ROSEVELT WAS.		No MUSPACE, THE ONLY DEMOCRATS WHO ONLY UN HAVE A MUSPACHE WIFI NO BEARD. WILSON HAD NEITHER. [1936] NO PRESIDENT'S BEEN		
IND GEN VELETICE	ROOSEVELT WAS. ROOSEVELT WAS. IT PAR ND ONE WHO GOT TEN MILLON VOTES HAS LOST.	MILLINARY HAS WON.	No MUSTRAE, THE ONLY DEMOCRATS WHO CAN UN HAVE A MUSTRAE UITH NO BEARD. WILSON HAD NEITHER. NO PRESIDENT'S BEEN REFLECTED WIT DUBLE DUGT UNEYROMENT.		NO DEMOCRAT HAS WON DURING WARTING.
IN DUAN NEEDEN DE UMTIL MCKIMEY WAS IN ONE WITH TWO CS IN THEIR NAME HAG DECOME PRESIDENT	III ROOSE/FELT WAS.	MILLINKY HAS WON.	No MUSRAE, THE ONLY DEMOGRITS WHO GAN WITHING NO SEARCH. WITHING BEARD. WILSON HAD NETFER. NO PRESIDENT'S BEEN REFLECTED WITH DUSE- DIGT WERPONTENT.	VILSON DD. VILSON DD. ITHOU NO CHE HAGWON A THIRD TERM.	UNTIL HARDING.
OMIT MELLEN VALLEN VAL	MONEVELT LAS.	MILLINKY HAS WON.	No MUSROE, FIE ONLY DENOSRIS UHO ON UM HAVE ANUSROE UIFH NO BERRD. , ULSON HAD NETFER. , ULSON HAD NETFER.	VILSON DD. VILSON DD. NO ONE HOS WON A TRIKO TERM.	UNTIL HARDING.
IND CENTRELECTER INDIALE MICHINE WAS INDIALE MICHINE AND BECOME PRESEDENT INTELECTUNIC RECORDER. INTELECTUNIC RECORDER.	Roode/ell unc.		No Phuspare, Fie Only Denocrits und orn win Have Anystrote With No Berro. Wilson Had Netfer. No President's Been No President's Been Contraction Durits Dest'und Photose Dest'und Photose Des		UNTIL HARDING.
		HUNRY HIS UON.		WISON DO. WISON DO. WISON DO. WISON DO. WITH PROVING WITH FOR DO. WITH FOR DO. WITH FOR DO.	UNTIL HARDING.
IND VEN VELEVILE IND OLE WHET WO CS IND ONE WHET WO CS IN THEIR WHET HAS BECOME PRESEDENT INTER VENUE COURSE INTER VENUE INTER VENUE INTE		HUTHAY HIS JON.	NOMESTICE THE OWN DENORING THE OWN UNITHER PUSTICE UNITH NO REFERENCE NO PRESIDENT SEED. NO PRESIDENT SEED. NO PRESIDENT SEED. ORTOUCH COMMENT.		NUMIL HARDING.
IND DEN NEELS UE	INDOREVET LAPS		NOMESTICAL ELECTRY DENORMES LUDOW WI HAR AND SERVED. ULSON HAD NETFER. NO PRESIDENTS BEEN NO PRESIDENTS BEEN NO PRESIDENTS BEEN CONFOLICIÓN DUBLECTRY NO PRESIDENTS DE SEN LOUTEL FDR LUAS.	UNISON DO. UNISON DO. NO ORE HIS UNIA A THRO TERM. UNITL FOR DO. EVERY REPORTION WOR FRANCOUSINA HIS LON.	UNTIL HARDING.
IND OLEN VIELES		HUTHAY HIS JON.	NOMESTICE THE OWN DEMORTS LUD OWN UN HING PUTSTORE UNTIL NO REPRO- NO PRESIDENTS BEEN NO PRESIDENTS BEEN REPRESIDENTS BEEN CONFIDENCE OWN HOUSE DOST DUEPROMENT	UNISON DD. UNISON DD. UNIC HIGUNN A THIRD TERT. UNIC HDR DD. EVERY REPORTED HER (DUSING HIG UDN)	UNTIL HARDING. IPPRIM ND DEPORATIONS ND DEPORATIONS UNITIONS WARTINE. UNTIC PDR DIR IPPRIM MO REPUBLICAN VICE RESOLATION VICE RE
IND DEN NEELEVILE INDIAL MEKINLEY WAS INDIALE WIFI TWO CS INTERINATE HAS DECOME PRESEDENT INTERINATE AND DEMORTHS CAN'T UN WIFLOUT ALABAMA.	ROSEVELT LAS. ROSEVELT LAS. NO DRE WHO GOT TEN MULLON IOTES HAG LOSE: ONTIC AL SMITH. ONTIC AL SMITH. ONTIC AL SMITH. 		WHUSDATE THE OWN DEMORPHS WHO OWN UN HIVE A MUSTACE WHIT NO BEPRO. UNISON HAD NETFER. UNISON HAD NETFER. DRESON THE BED. NO PRESIDENT SEED. NO PRESIDENT SEED. UNISON HAD NETFER. UNISON HAD NETFER. UNISON HAD NETFER. UNISON HAD NETFER. UNISON HAD NETFER. UNISON HAD NETFER. UNISON HAD NETFER.	UNISON DO. UNISON DO. UNITE PREMON A THEO TEAM. UNITE FOR DO. LIGHT DEAY REPUBLICAN WOST REPUBLICAN HIG LON.	UNTL HARDNG.
I THO DEAN NEEDENTER ONTIC MCKIMEY WAS NO DAE WITH TWO CS NT THEIR NAME HAS DECOME PRESIDENT UNTIC ONUN COLDEC: UNTIC ONUN COLDEC: UNTIC ONUN COLDEC: 			NO-MUSTRICE, THE OWN DEPORTING THE OWN WITH IN HUTCH CHARGE WITH IN DEPORT UNITING DEPORT UNITING THE OWNER DEPORT DEPORTUGE DEFORMED WITH DURGE DEFORMED WITH DURGE DEFORMED WITH DURGE DEFORMED WITH DURGE OWNER DURGE OF OWNER OWNER DURGE OF OWNER UNITING FEMALED WITH DURGE OWNER DURGE OF OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER DURGE OWNER OWNER DURGE OWNER DURGE OWNER OWNER DURGE OWNER DURGE	UNISON DR. UNISON DR. NO CHE HYSURUN A THRO TERT. UNITL FOR DR. LICEY TERT. UNITL FOR DR. UNITL GOLDWATER UNITL GOLDWATER	UNTIL HARDING.
IND OLEN NELLSING IND OLE WIRT TWO CS IND ONE WIRT TWO CS INTERINVE HAS BECOME PRESDENT INTERINVE HAS DECOME PRESDENT INTERINVE HAS INTERINVE HAS INTERINVE INTERINVE HAS INTERINVE HAS INTERINVE INTERINVE INTERINVE INTERINVE INTERINVE INTERINVE INTERIN INTER	III TOUSEVELT LAPS. III TOUSE LAPS CONTINUES HAP LOCAL VOCAT IIII TOUSE LAPS CONTINUES IIIIII TOUSE LAPS CONTINUES IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII		WHISTING TE THE ONLY DEMORPHS WHO CHILD UN HIVE A MUSTICALE WHIT NO BEPRO. UNISON HAD NETFER. UNISON HAD N	WILSON DO. WILSON DO. WILL POR DO. WITL FOR DO. DERY REPUBLICAN WICH THEN LOUGHAN HIS LON. 	
UNITLORUM CAUDEL UNITLORUM CAUDEL UNITLORUM CAUDEL UNITLORUM CAUDEL UNITLORUM CAUDEL UNITLORUM CAUDEL INTRUMM CAUDEL TRUMM DA TRUMM DA TRUMM DA TRUMM DA TRUMM DA			MULTING CONTRACTOR CON	UNISON DO. UNISON DO. NO OKE HYSUNN A THRO TERM. UNITIL FOR DO. LIBERT LIBERT LIBERT LIBERT LIBERT MONE LIMEN COUSAWA HIG LONNIC COUDWATER IMONE WITH TWO MODEL NIMES HIG GEODYE RESIDENT	
UNITL MCKIMEY WAS UNITL MCKIMEY WAS UNITLOUND AND ME WIFE HAS BECOME PRESIDENT. UNITLOUND COLORCE. UNITLOUND COLORCE. DEMOCRIES CAN'T WI WIFFOUT ALBOMMA. TRUMAN DID. TRUMAN DID.	BONFUELT I.APS. ROWEVELT I.APS. NO ONE UNO GOT TEX MILLION UNDES HAS LOST. OWITE AL SMITH. INGEL		NONESTIGE TE ONY DENOTING TE ONY DENOTING UTO ON UNIT NO NETTO DENOTING UNIT NO NETTO DETAIL NO PRESIDENT BEDI RELECTED NO NETTO DETAIL NO RESIDENT MON RELECTED NO NETTO DENOTING ONFOLIS CAN'T UM.	UNISON DO. UNISON DO. UNITIC DOR DOR UNITIC FOR DOR UNITIC FOR DOR UNITIC COCLUMPER UNITIC COCLUMPER UNI	
IND DEAN NEEDEN DE IND DEAN NEEDEN DE IND ONE WITH TWO CS INTHEIR NAME AND BECOME PRESEDENT INTHE NAME AND DEPORTS ONT UN WITHOUT AURONAL DEPORTS ONT UN WITHOUT AURONAL INTER NAME INTER N	ROUGEVELT LASS     TO THE WHO GOT     TO THE HOLE AL SPITIFL     TO THE HOLE AL SPITIFL     TO THE WHO COT	HUTHAY HIS JON.		UNISON DO. UNISON DO. ITHO: NO OKE HYGUNA A THRO TERM. UNITE FOR DO. DERY REPLEMENT MOS REPLEMENT INFORMATION NO OKE UNITE FOR NO OKE UNITE FOR MOS PROJOCIAL MOS PRO	
IND DEAN RELEASE	BOSEVELT LAPS. BOSEVELT LAPS. NO DRE WHO GOT TEN PILLON VOTES HAFG LOST. ONTIC AL SMITH. DISCHOLER DR. BISCHOLER DR. BUSCHOLER DR. BUC CARFER DR. BUC CARFER DR. 	HUTHAY HIS UON.	NO MUSTINGE THE OWN DEPORTING WID OWNER WIFFI NO DEPORD WIFFI NO DEPORD UNESN HYDO NETFERE UNESN HYDO NETFERE UNESN HYDO NETFERE UNTEL FDR WAS UNTEL FDR WAS UNTEL FDR WAS UNTEL FDR WAS UNTEL FDR WAS UNTEL REPORT WAS DED REDUCTED WESDENT HYDO DED REDUCTED WID DATE DED REPORT HYDO	UNISON DDC. UNISON DDC. III OXE HYSUDW A THRO TERT. UNITL FDR DDC. III DENT. DUCK MEDICION WOST FRANCUCON WOST FRANCUCON MODEL MUTH THO FRANCUCON NO OKE LIMIT THOSE BECOME PRESONT. UNITL'HERBERT LIMUER'	
IND DEAN NEEDENTER UNITE MEKINDE Y WAS NO DUE WITH TWO CS NI THEIR NAME HAG BECOME PRESDENT UNITE CHUNN (COLDICE. 1998) DEMOGRATIS CAN'T UN WITHOUT ALAGAMA. UNITE NIXAN DO. 19752. CAN'T WIN TWOCE. UNITE NIXAN DO. 19752. UNITE NIXAN DO. 19752. UNITE NIXAN DO.		HUTHAY HIS JON. UNTIL THAT: UNTIL THAT: UNTIL THAT: UNTIL THAT: UNTIL THAT: UNTIL THAT: UNTIL EXCAN EAST SCOND THE IN A SCOND TH	NOMESTIGE TE ONE DEPORTS LIDON UNITAL PLANTAGE UNITAL DEPROL NO PRESIDENT SEEN RELEATED VIEW HOMETHER UNITAL PERMEMAN ORFOLISS CANT UNIT CATFOLISS CANT UNIT CATFOLISS CANT UNIT CATFOLISS CANT UNIT CATFOLISS CANT UNIT CATFOLISS CANT UNIT CATFOLISS CANT UNIT DE DETFINADED PRESIDENT INS BEEN REBEATED. UNITAL REPORT UNITS CATFOLISS CANT UNITS CATFOLISS CAT UNITS CATFOLISS CAT UNITS CATFOLISS CATT UNITS	UNISON DO. UNISON DO. UNITU FOR HOLON A THRO TER1. UNITU FOR DOL. UNITU FOR DOLLAR DUCKY REPUBLICAN WOO THRO LONG UNITU GOLDWATER UNITU FOR DOLLAR UNITU FOR DOLLAR UN	
UNTEL MERINE HAS BECOME PRESIDENT. DI DOLE UNE TIMO CS IN THEIR MARINE HAS BECOME PRESIDENT. DEMORRIS CANT UN WIFFOUT ALABAMA. UNITE NUMBER CANT WIN TIMICS. UNITEL NUMBER CANT WIN TIMICS. UNITEL NUMBER COMERS. CANT WIN TIMICS. UNITEL NUMBER COMERS.	BONEVELT I.APS. REVEVELT I.APS. NO ONE UNO GOT TEM MULLON UNDES HAS LOST. DATIF, RUSPITIF. INGEVENT UNDES BORNOUER DIR BORNOUER DIR BORNOUER DIR 		NOMESTIGE TREAMS DEPOSITE UND ON UNTERNISTICE UNTERNISTICE UNTERNISTICE UNTERNISTICE UND RESIDENT BEDI RESELECTED UND REFER UNTER FOR WAS. UNTER FOR WAS. UNTER RESOLUTION ON DEPT-INVOCED PRESIDENT IMS BEDI RESELETED. UNTERREAM UND. UNTERREAM UND.	I UNISON DO. I UN	
INFOLDENT RELEATER     INFOLDENT RELEATER     INFOLDENTE HAG     BECOME PRESIDENT     INTERENTATION COUDCE.     INFOLDENTE ANG     BECOME PRESIDENT     INFOLDENTE ANG     BECOME PRESIDENT     INFOLDENTE ANG     INFOLDENTE ANG     INFOLDENTE ANG     INFOLDENTE ANG     INFOLDENTE ANG     INFOLDENTE ANG     INFOLDENTE		HUTHAY HIS JON. UNITU THAT. ITT22. ND DETOCRIT HAS WOR SKILL WHEN SEURED HE REAT TO VOTE. UNITU FOR DOL ITT22. UNITU FOR DOL ITT22. UNITU FOR REATION DOLE CAN DEAT HAS SALE JONE AN DOLE CAN DEAT HAS SALE JONE AN LOGY FOR REATION NO CAN DEAT HAS BEEN DOLE (IN DEAT HAS SALE JONE AN DOLE (IN DEAT HAS SALE JONE AN INTO IT CREASE HAS UNITU REPORT WAS UNITU REPORT WAS UNITURE REPORT		UNISON DO. UNISON DO. UNITE FOR HIGHNAN A THRO TERM. UNITE FOR DO. UNITE COLUMPER HERY REPLACION MOD RE UNITE COLUMPER UNITE COLUMPER	
IND DEAN NEEDEN DE     IND DEAN NEEDEN DE     IND DEAN NEEDEN DE     IND DAE WIRF TWO CS     INTERENANTE PHO BECOME PRESEDENT     INDIA     INTERENANTE PHO INDIA     INTERENANTE     INDIA     INTERENANTE     INDIA     INTERENANTE     INDIA     INTERENANTE     INDIA     INTERENANTE     INDIA		HUTHAY HIS JON.	NUMERAL REAL AND A CONTROL OF THE CO	UNISON DO. UNISON DO. UNITE FOR HOLMAN A THRO TERM. UNITE FOR DD. UNITE FOR DD. UNITE COLUMPER MOD FROM COLUMPER UNITE COLUMPER MOD FROM COLUMPER MOD FRO	UNTIL HARDINGUNTIL HARDINGUNTIL POR DIRUNTIL FOR DIRUNTIL FOR DIRUNTIL FOR DIRUNTIL FOR DIRUNTIL NIKONUNTIL NIKONUNTIL NIKONUNTIL NIKONUNTIL CLINTON DIRUNTIL VIGGE
IND DEAN RELEASE		HUINAY HIS JON. UNIL TART IND REPORT HIS VOI SAULUTE IND REPORT HIS VOI SAULUTE IND REPORT HIS SAURD HE RAFT IND REPORT IND RE	Nonusingle The Own Deroging with own Units of the Own With Wein House Units No Derfoo Units No Derfoo Units of the Owner Belleted Durits Beah Conforder Content Onerower Conforder Content Onerower Conforder Content Onerower Conforder Content Owner Der Content Conforder Content Content Conforder Content Conforder Content Content Conforder Content Content Conforder Content Conforder Content	UNISON DO. UNISON DO. UNIC FOR DO. UNIC FOR DO. UNIC FOR DO. UNIC FOR DO. UNIC GOLDWATER HOL LIFT TO MODE LIFT TO MODE LIFT TO UNIC COLDWATER HOLDE LIFT TO UNIC PRESENT LIVER CHELONGERS UNIC PRESENT LIVER CHELONGERS UNIC PRESENT LIVER CHELONGERS UNIC PRESENT LIVER UNIC PRESENT LIVER	UNTIL HARDINGUNTIL HARDINGUNTIL HARDINGUNTIL FOR DIZUNTIL FOR DIZUNTIL FOR DIZUNTIL POR DIZUNTIL NIXONUNTIL NIXON.

Prediction functions that work on the training data might not work on other data

Minimizing the training error is a reasonable thing to do, but it's possible to minimize it "too well"

• If your function matches the training data well but is not learning general rules that will work for new data, this is called **overfitting** 



From: <u>https://www.quora.com/Whats-the-difference-between-overfitting-and-underfitting</u>

Restrictions on what a classifier can learn is called an **inductive bias** 

Inductive biases are an important (and actually necessary) ingredient to learning classifiers that will generalize to new data

One type of bias: don't use certain features

Age	Previous Purchase	Likes Kanye	Likes New TSwift
13	Ν	Ν	Υ
15	Ν	Υ	Ν
16	Ν	Ν	Υ
22	Ν	Υ	Ν
28	Υ	Ν	Υ

One type of bias: don't use certain features

Age	Electric Toothbrush	Likes Kanye	Likes New TSwift
13	Ν	Ν	Υ
15	Ν	Υ	Ν
16	Ν	Ν	Υ
22	Ν	Υ	Ν
28	Y	Ν	Y

We know from common sense that this is probably irrelevant, and any association is a coincidence

Another type of bias: restrict what kind of function you can learn



Linear functions (lines or planes) are so simple that they won't overfit, even if they aren't perfect on training data

We'll discuss other types of inductive bias (some automatic) that can help with generalization throughout the semester

#### Almost done

# Uncertainty

When making a prediction, there is some uncertainty (by definition)

Many machine learning models can estimate the **probability** that a particular prediction is correct

### Machine Learning in Practice

