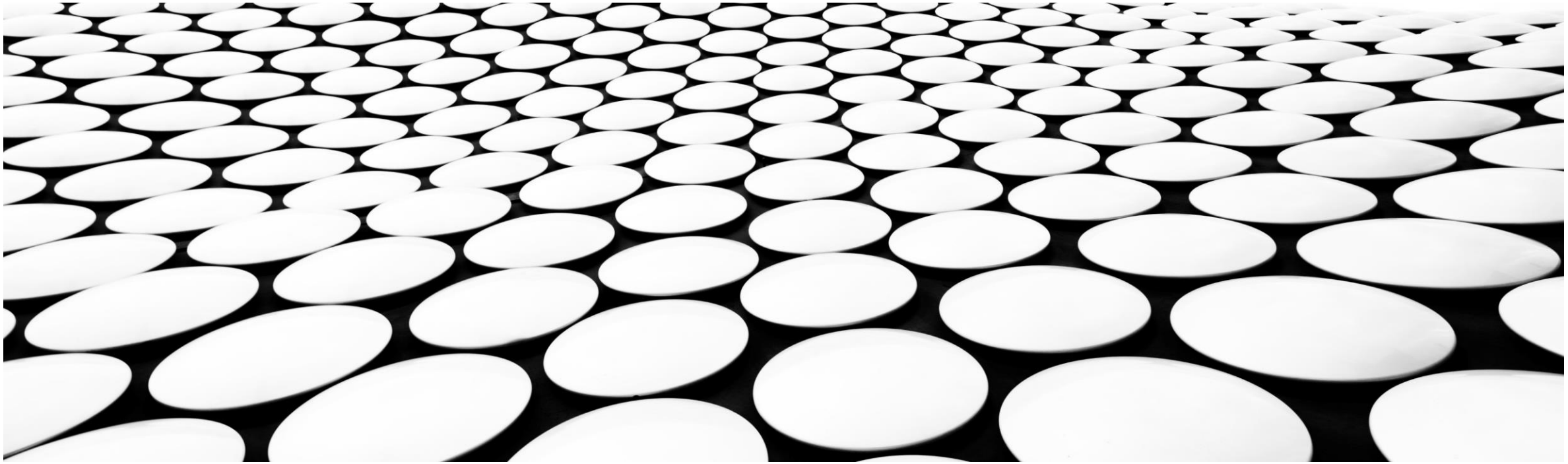
SUPERVISED MACHINE LEARNING

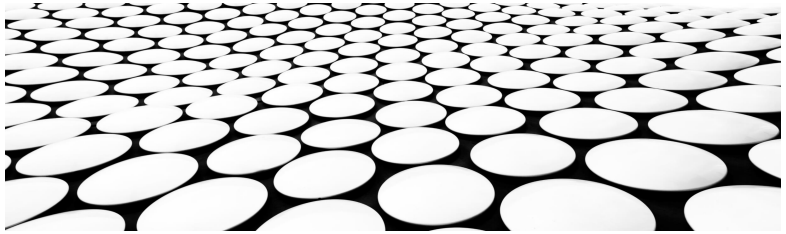
PART 3 OF 4 – GENERATIVE AND DISCRIMINATIVE LEARNERS

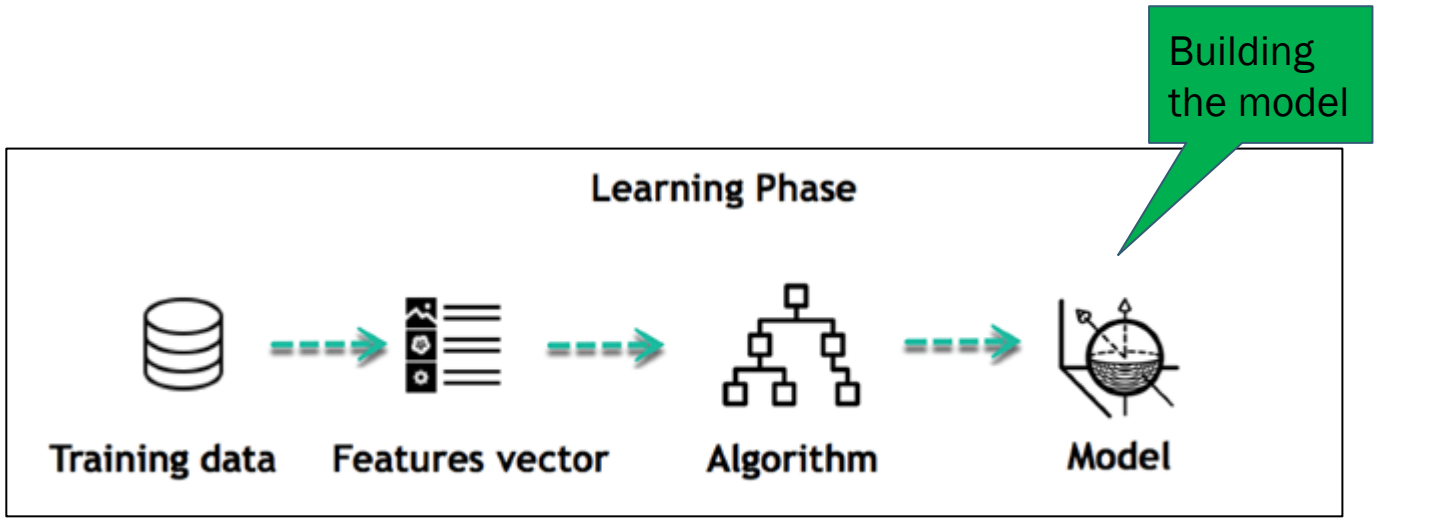


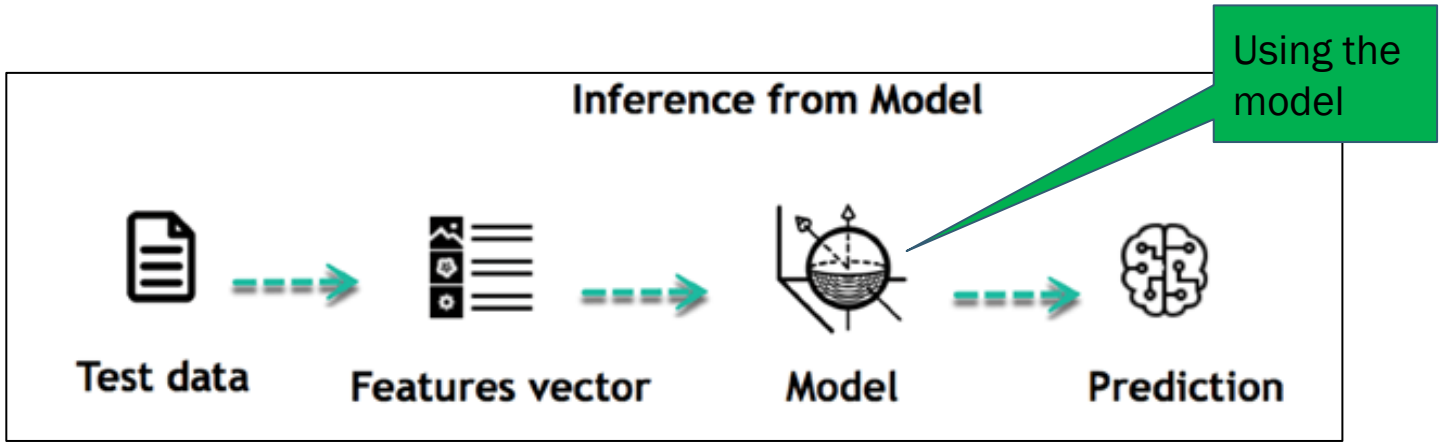
GOALS

- Present and compare:
  - discriminative learner
  - probabilistic learner

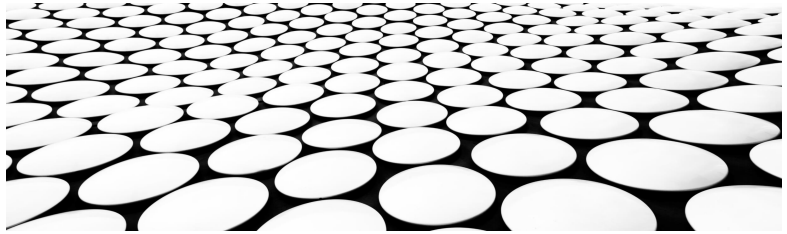
LEARNING / TESTING

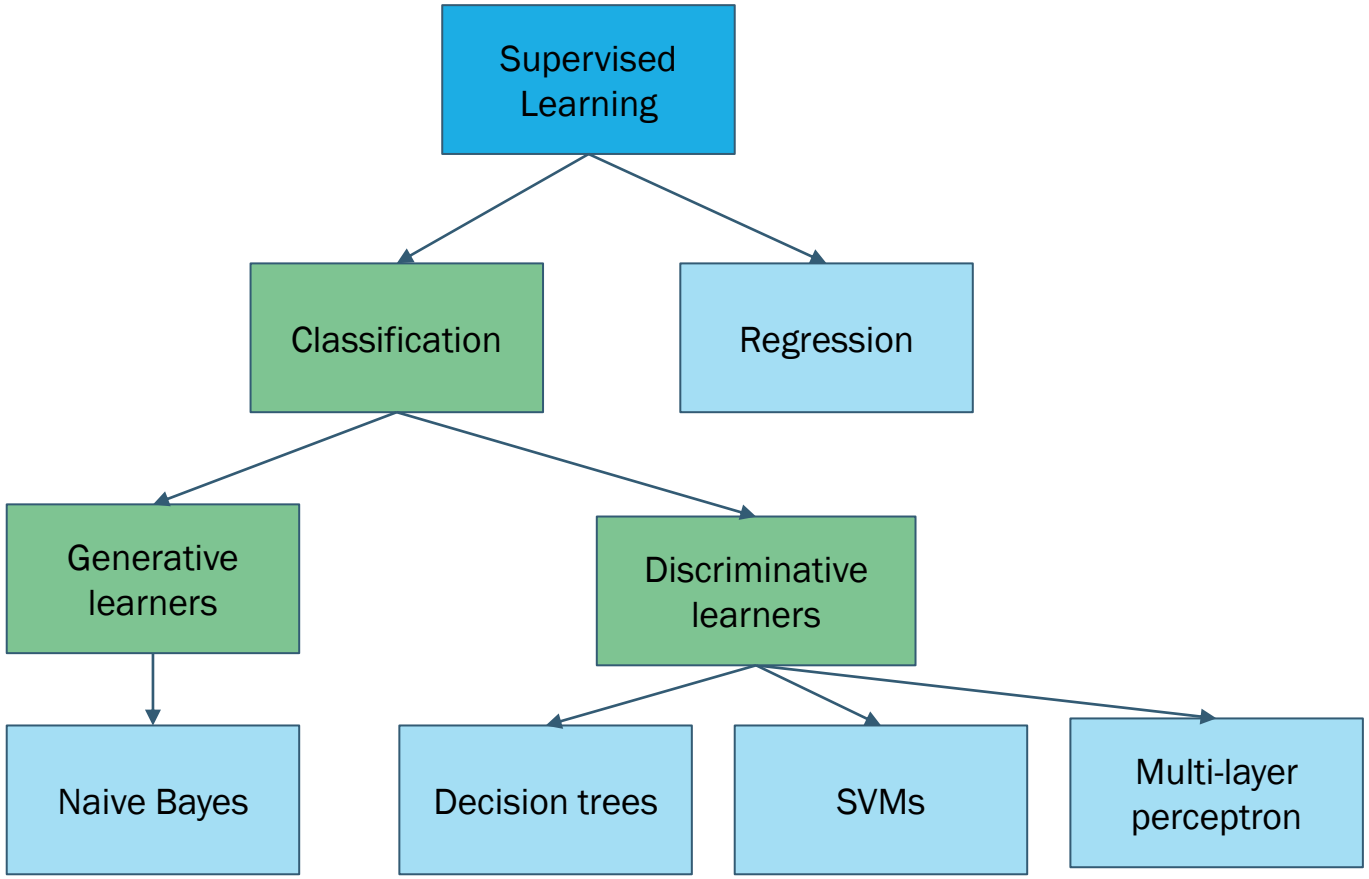




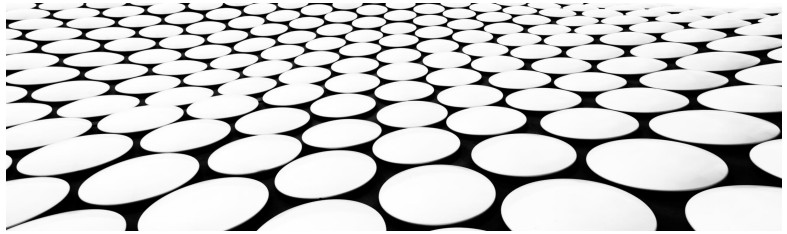


VARIOUS MODELS





DISCRIMINATIVE LEARNER



Objective of the learning algorithm

Learn a **decision surface** in an N-dimensional space (N - number of features) that distinctly classifies the data points (examples).

We call the decision surface, the discriminative function.

Generalization capability

Be able to correctly classify unseen examples.

BIKING/DRIVING EXAMPLE

Sample	Temperature	Rain/snow	Bike / Drive
s1	-22	10	Drive
s2	-24	12	Drive
s3	-15	20	Drive
s4	-8	13	Drive
s5	-5	40	Drive
s6	8	45	Drive
s7	-2	5	Bike
s8	5	5	Bike
s9	12	2	Bike
s10	8	5	Bike
s11	12	20	Bike
s12	13	15	Bike
s13	15	10	Bike
s14	18	5	Bike
s15	19	20	Bike

Training data

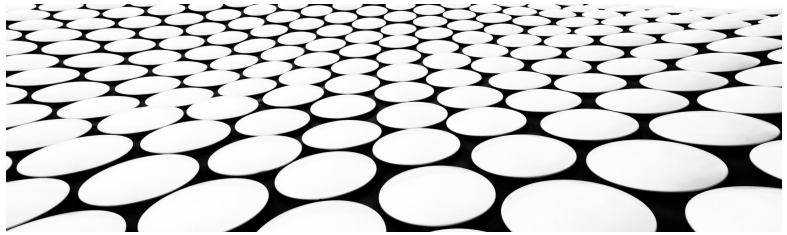
15 samples, s_i for i = 1..15

Feature vectors

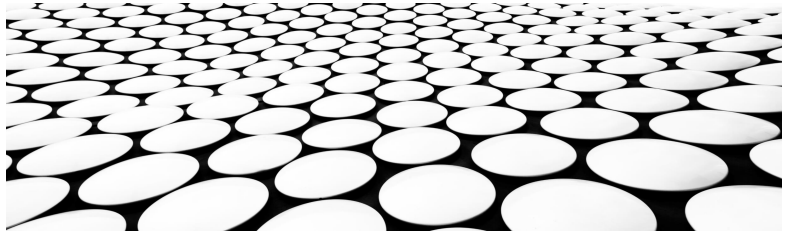
$\vec{x}$ has 2 dimensions

x_1 Temperature

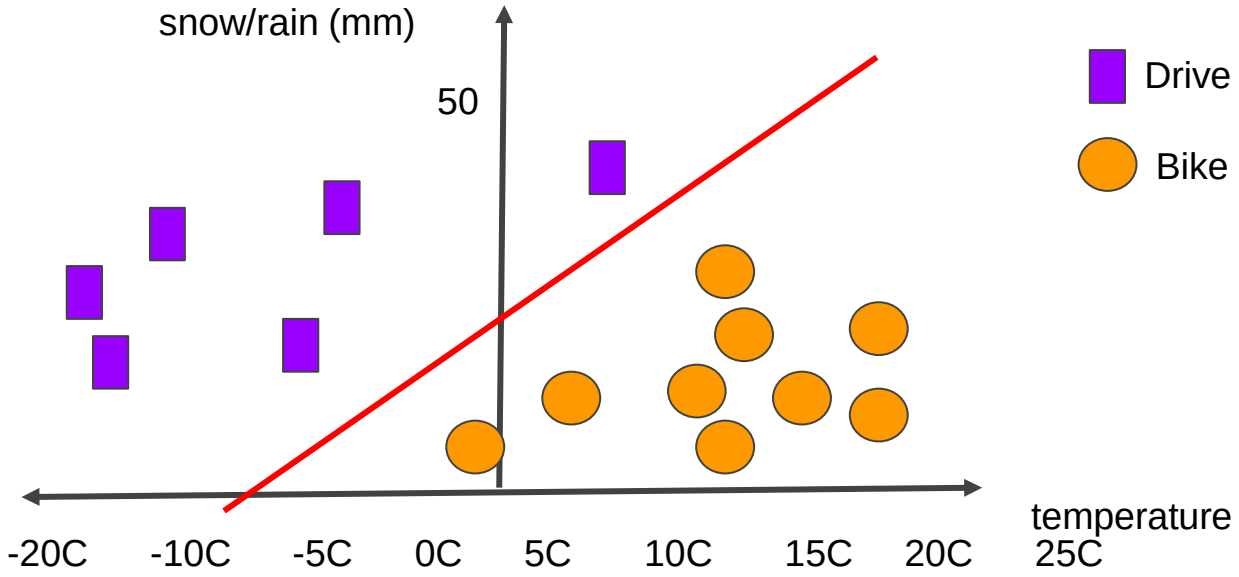
x_2 Rain/snow



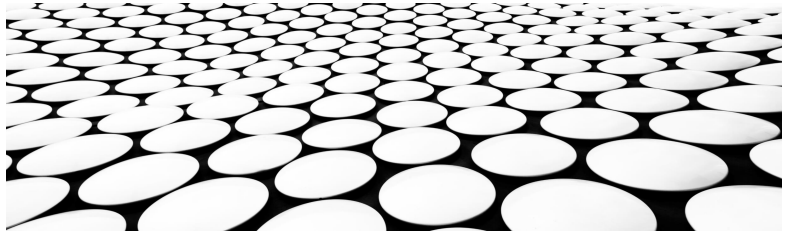
DISCRIMINATIVE LEARNER



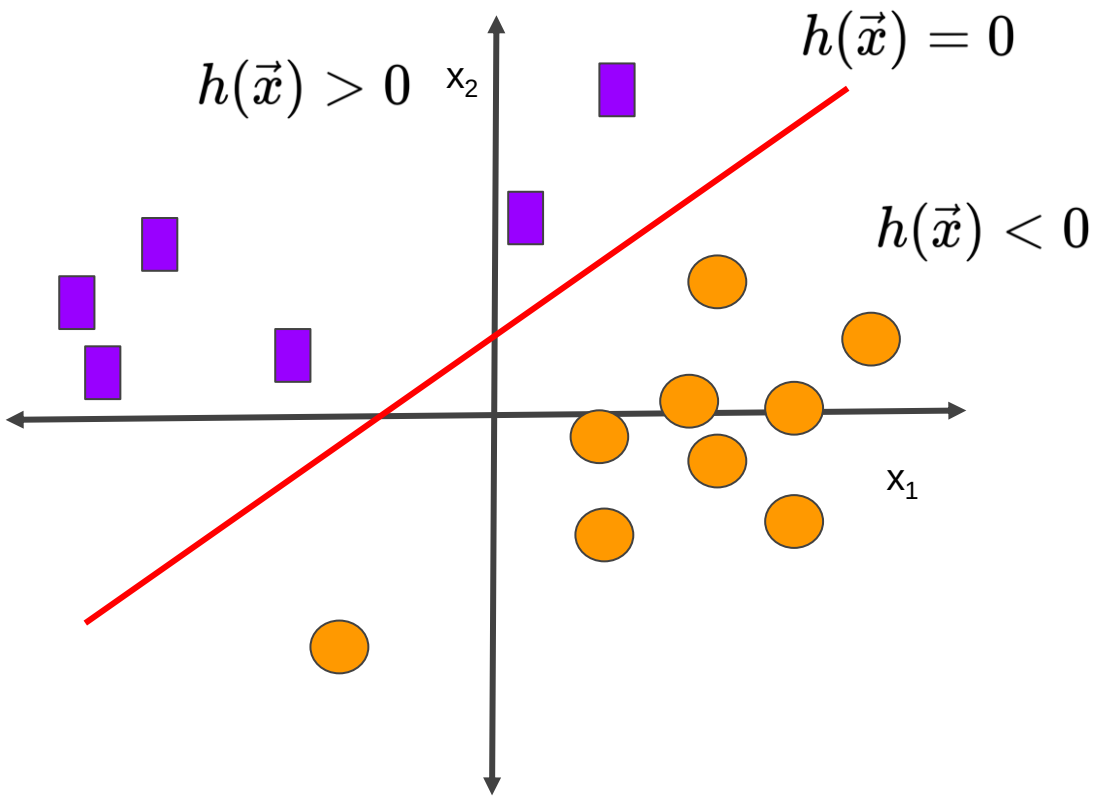
Find a decision surface (red line) that separates the training data.



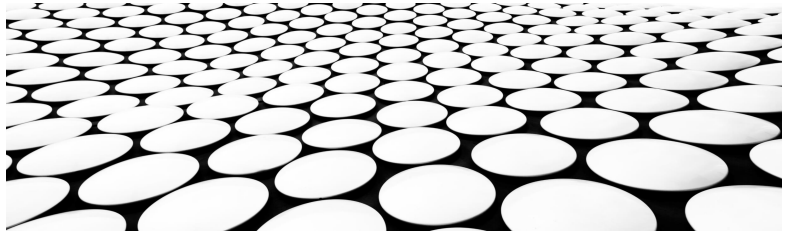
DISCRIMINATIVE LEARNER



$$h(\vec{x}) = \vec{w}^T \vec{x} + w_0$$



DISCRIMINATIVE LEARNER



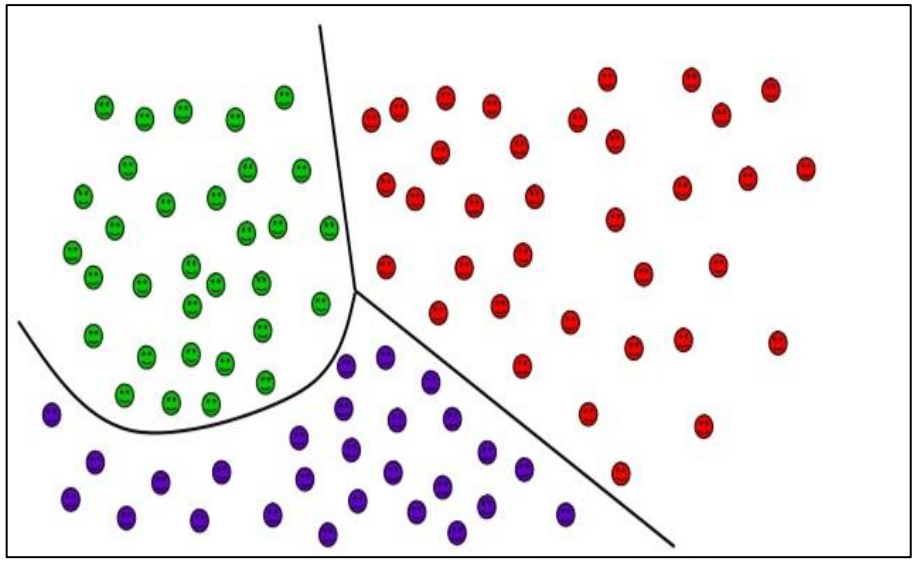
Multi-class

There can be multiple target classes.

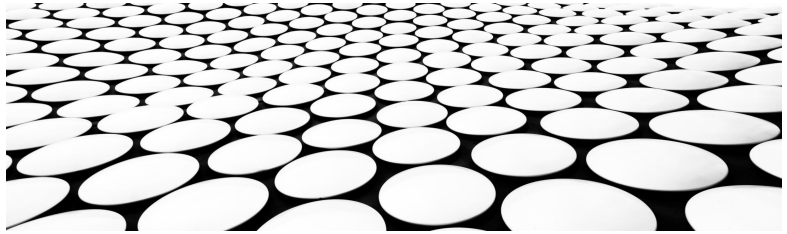
Various types of boundaries

Decision regions are shown (green, red or blue)

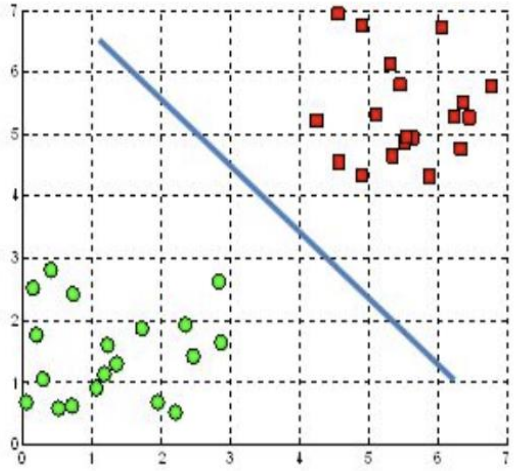
Decision surfaces are the frontiers between the decision regions.



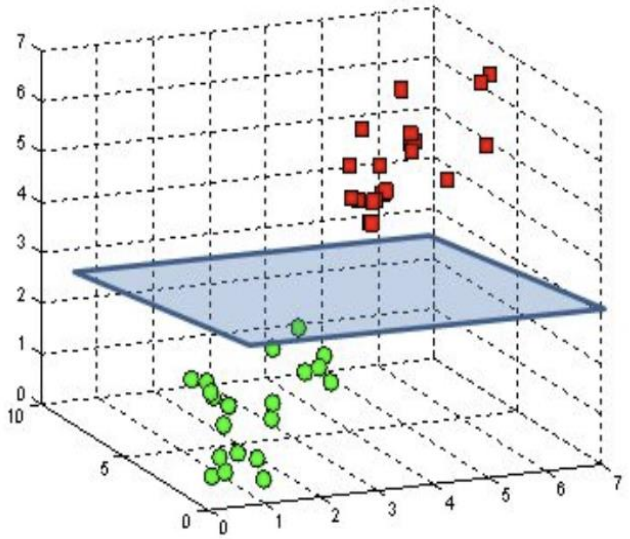
DISCRIMINATIVE LEARNER



A hyperplane in $\mathbb{R}^2$ is a line

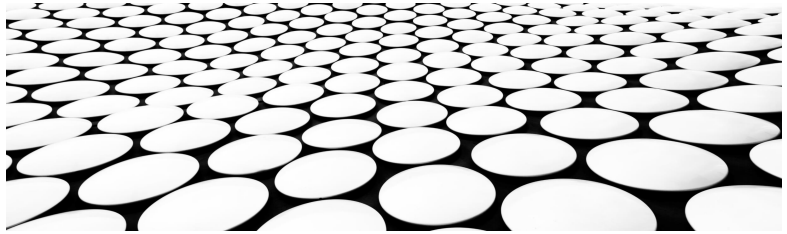


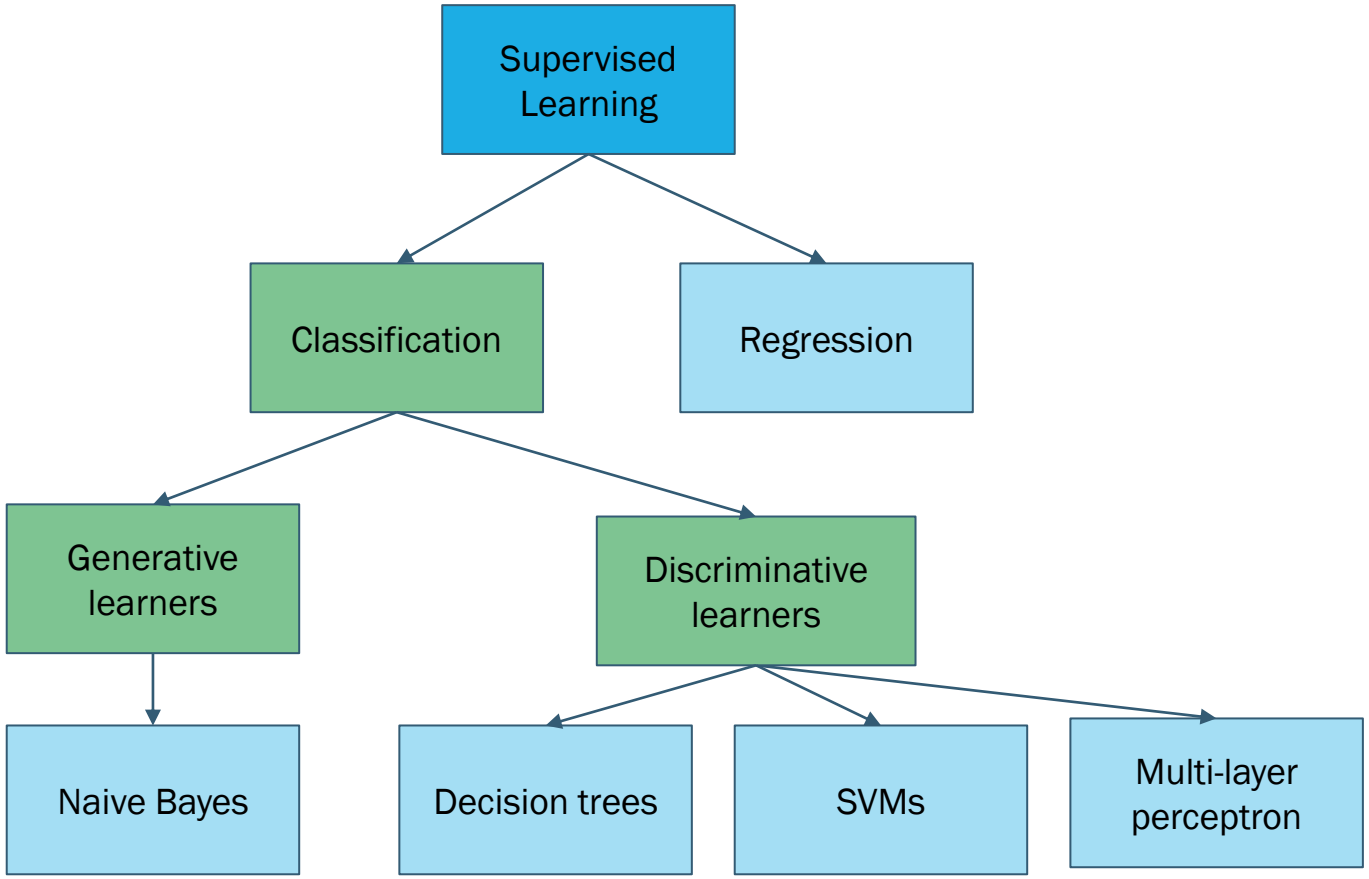
A hyperplane in $\mathbb{R}^3$ is a plane



Source

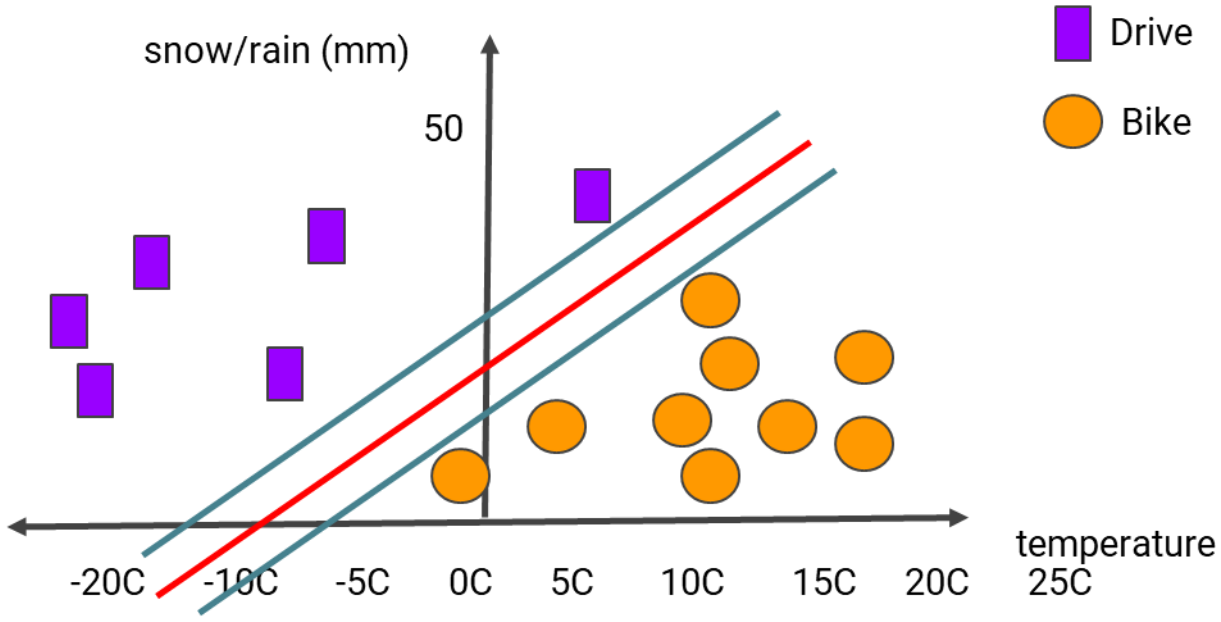
VARIOUS MODELS

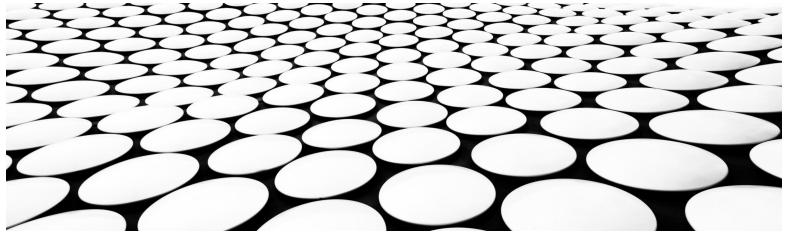




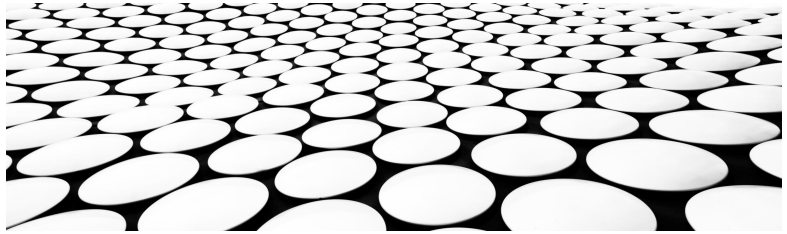
SVM AS DISCRIMINATIVE LEARNER

Maximize the margin between the support vectors

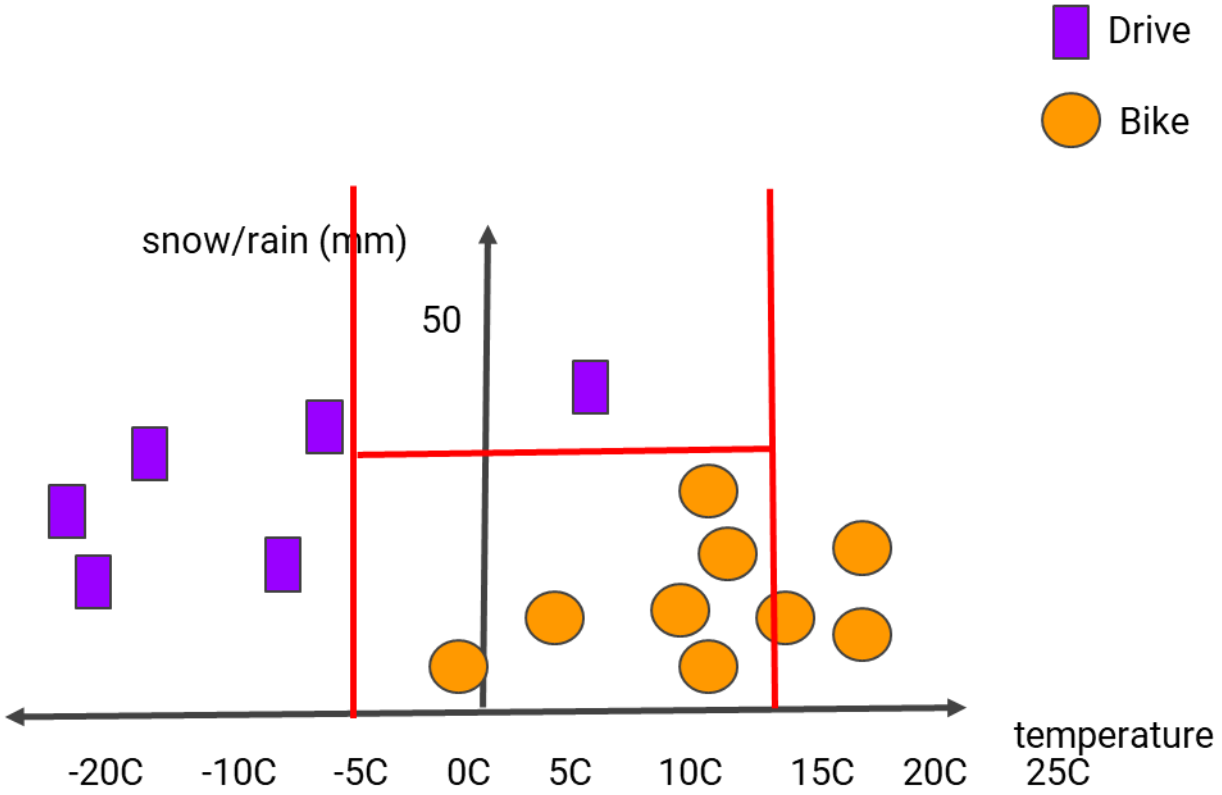




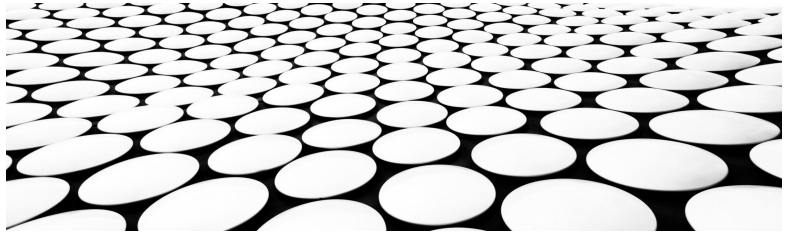
DECISION TREE AS DISCRIMINATIVE LEARNER

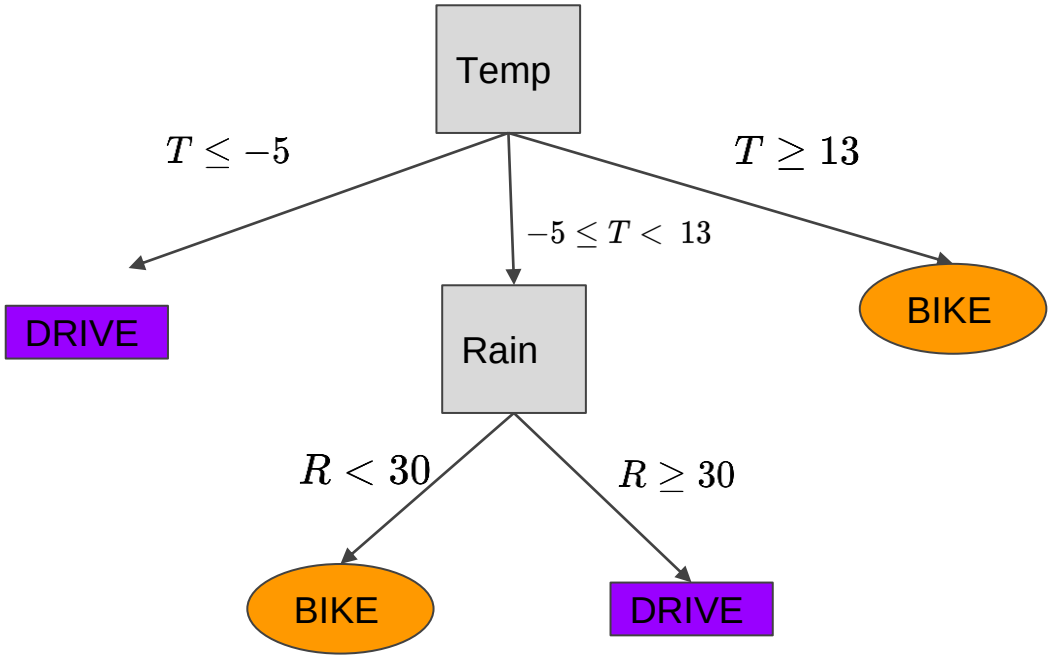


Will create as many boundaries as necessary to separate the training data.

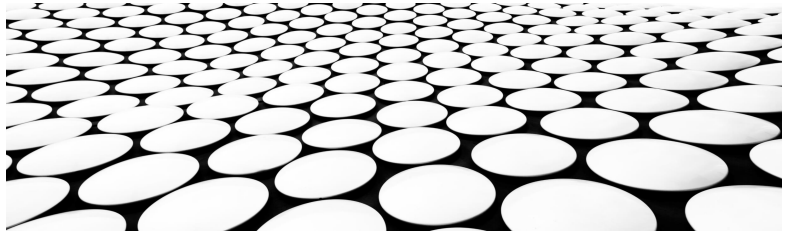


DECISION TREE AS DISCRIMINATIVE LEARNER

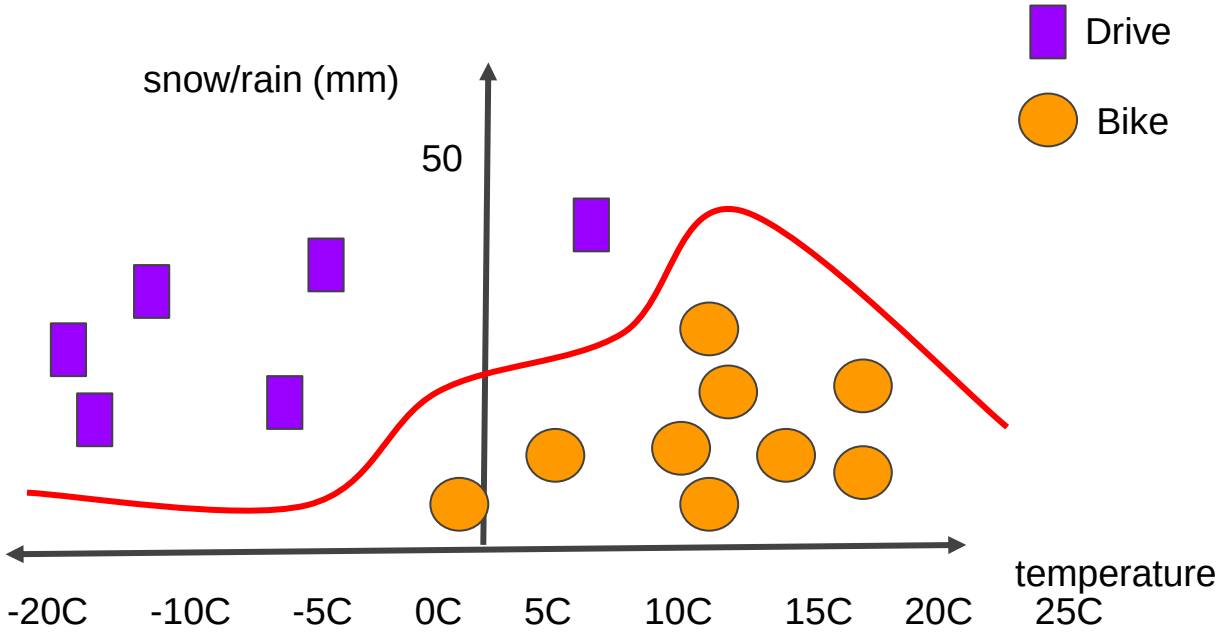




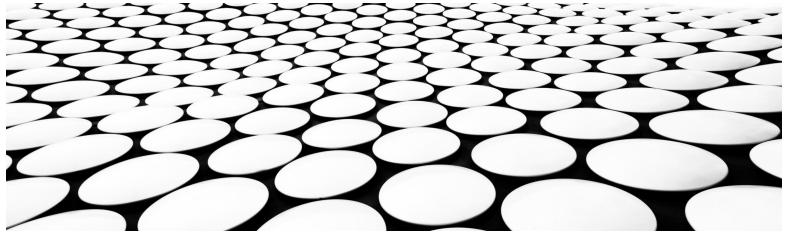
MULTI-LAYER PERCEPTRON AS DISCRIMINATIVE LEARNER

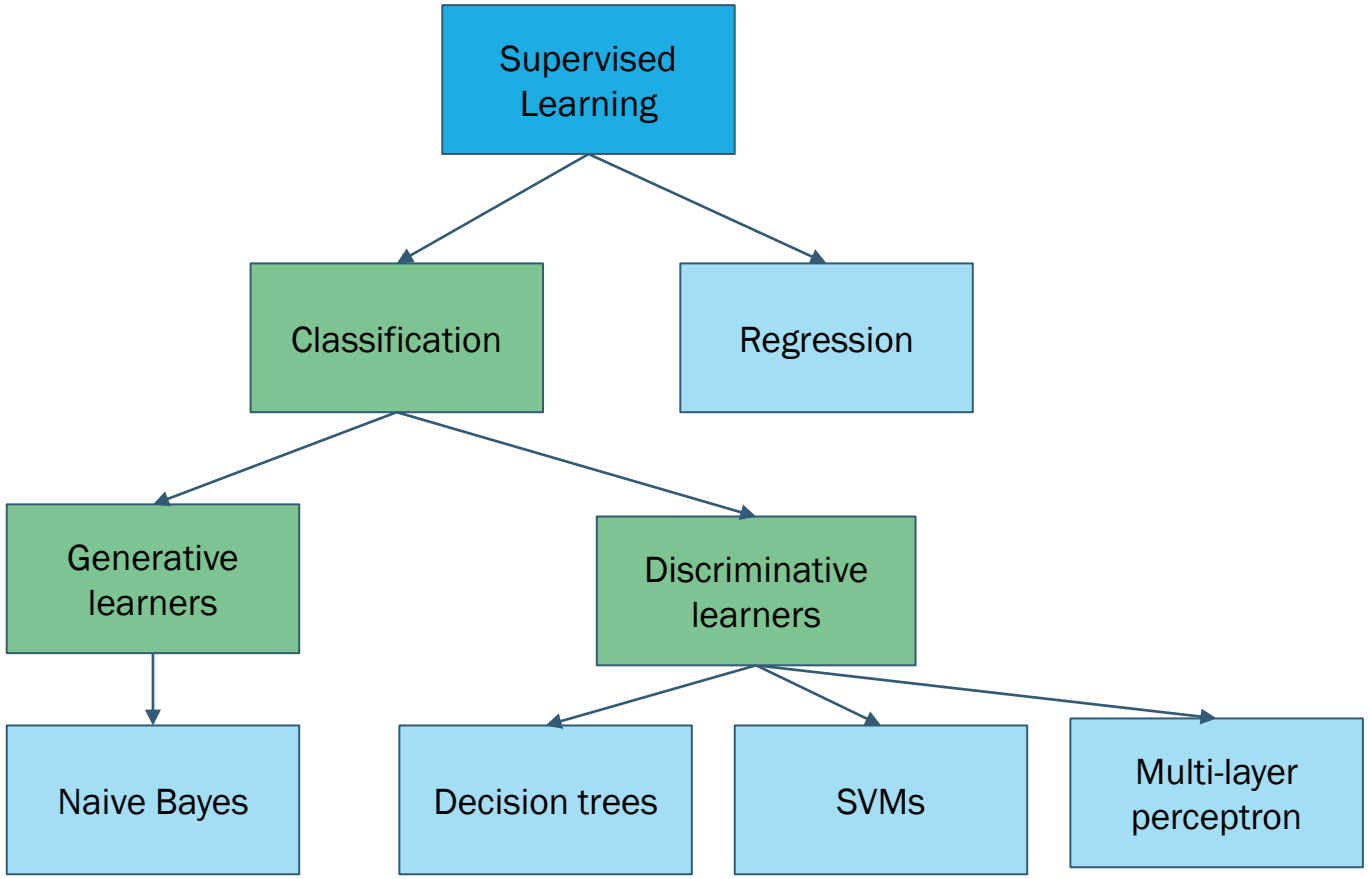


Will be able to learn decision surfaces that are not planes



VARIOUS MODELS

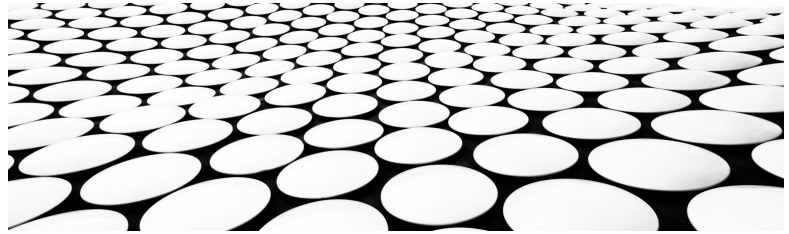


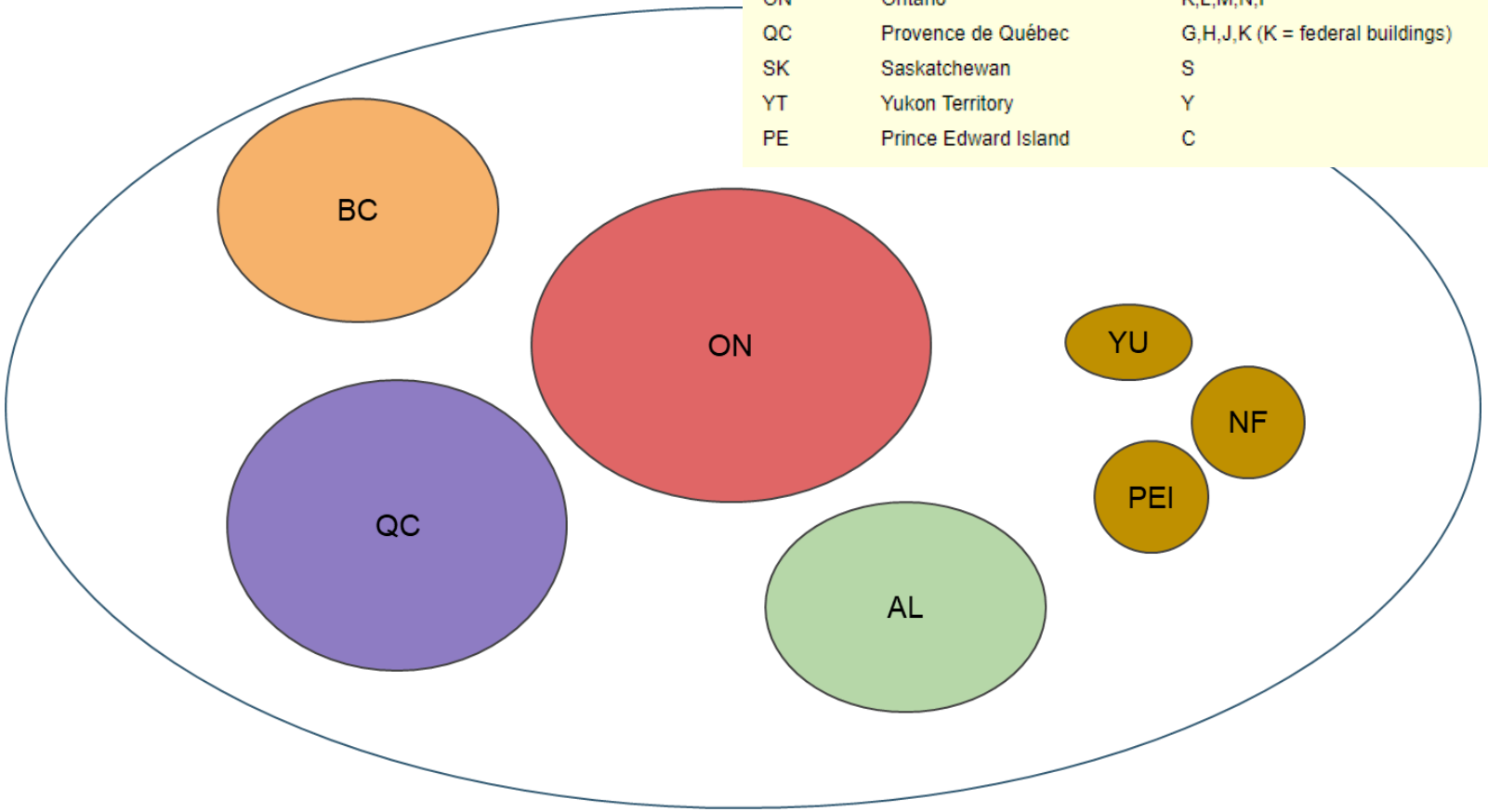


PROBABILISTIC APPROACH

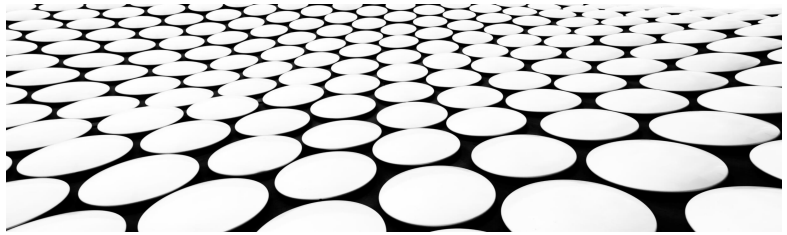
The learner assumes the data comes from more or less probable classes.

Code	Province name	First letter of postal code
AB	Alberta	T
BC	British Columbia	V
MB	Manitoba	R
NB	New Brunswick	E
NL	Newfoundland & Labrador	A
NS	Nova Scotia	B
NT	Northwest Territories	X
NU	Nunavut	X
ON	Ontario	K,L,M,N,P
QC	Provence de Québec	G,H,J,K (K = federal buildings)
SK	Saskatchewan	S
YT	Yukon Territory	Y
PE	Prince Edward Island	C





BIAS

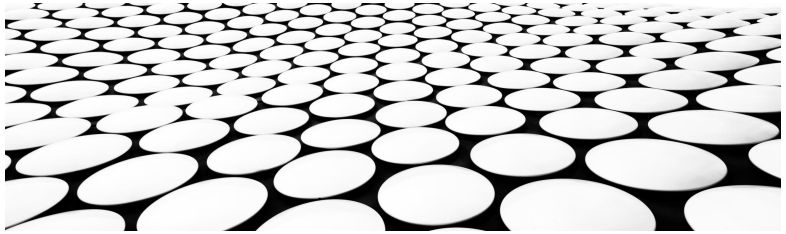


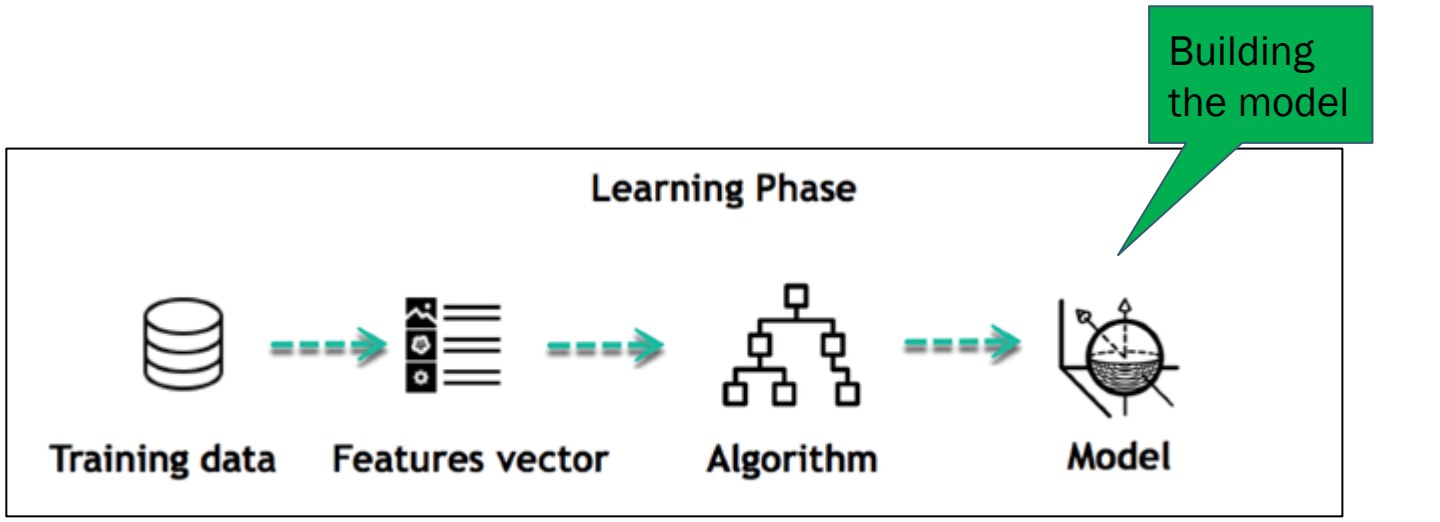
It is important that the construction of the training set reflects the generative process.

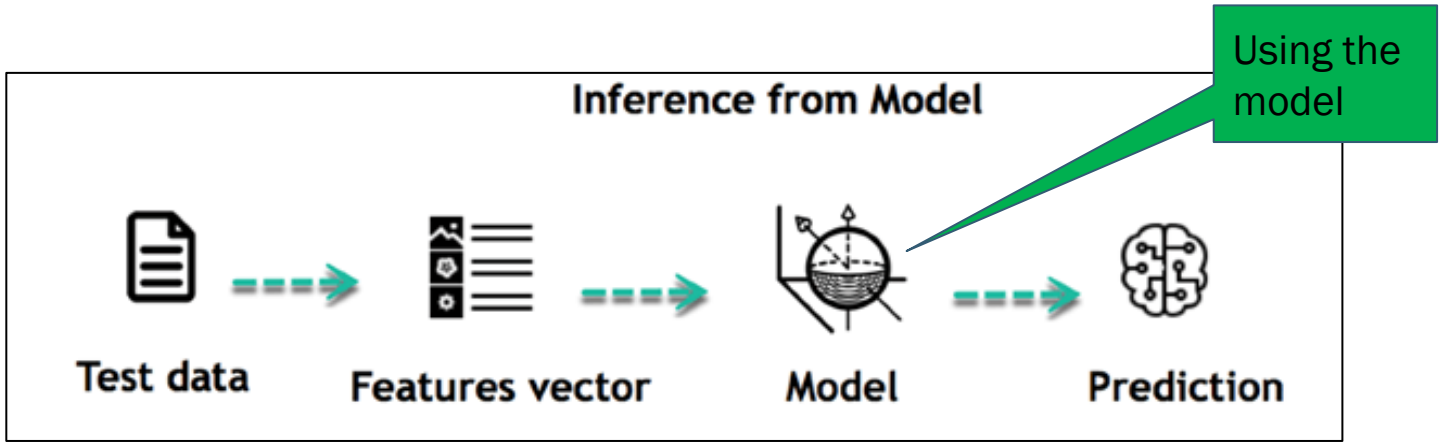
	Features			Possible classes	
Movie	Duration	Producer	Lead actor/actress	Genre	Viewer's appreciation
M1	90	P1	A1	Drama	Bad
M2	140	P3	A2	Drama	Bad
M3	80	P1	A1	Comedy	Bad
M4	90	P1	A3	Drama	Bad
M5	100	P3	A4	Drama	Bad
M6	110	P2	A5	Comedy	Bad
M7	100	P2	A5	Comedy	Good

Unbalanced training set

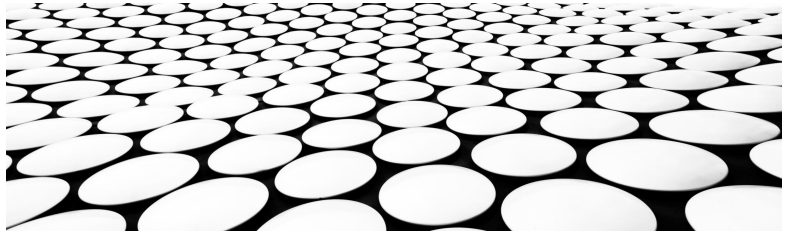
LEARNING / TESTING





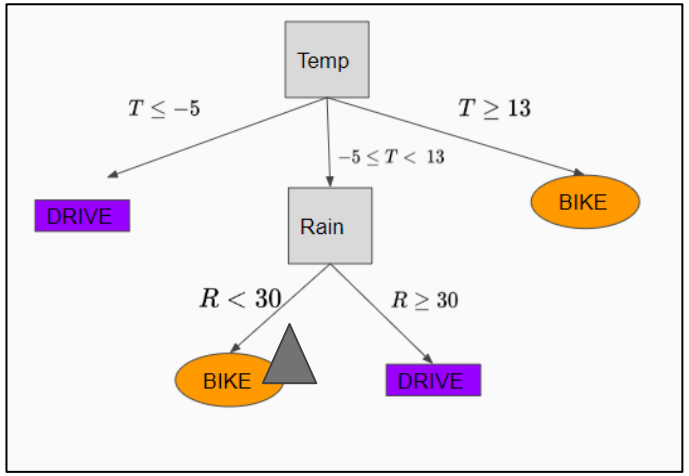


PREDICTION WITH DIFFERENT MODELS



Sample	Temperature	Rain/snow	Bike / Drive
T1	8	2	??
T2	15	20	??

Decision Tree



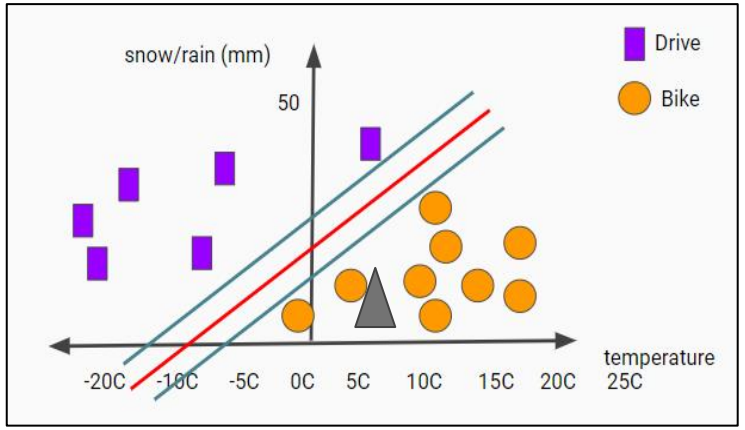
Naïve Bayes

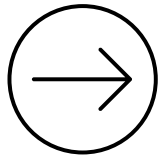
Hypothesis testing

P(Drive | T= Average & Rain = Normal)
P(Bike | T= Average & Rain = Normal)

Take max. --- Bike

SVM





LET'S CONTINUE…

Next video:
Evaluation