

# Machine Learning

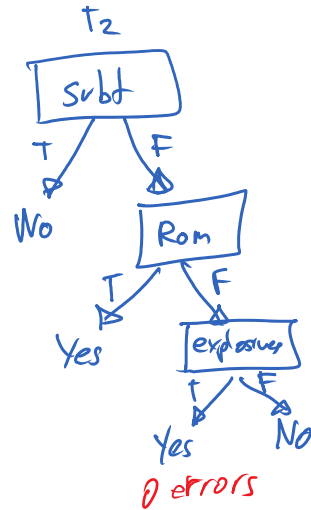
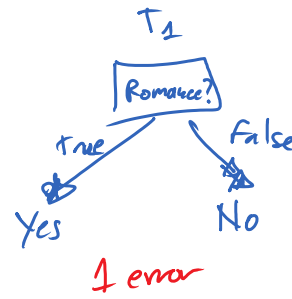
↳ Supervised Learning

↳ Basic Models — Decision Trees

— linear and non-linear classifiers

• Example: Movie Recommendation system

	$X_i$				$Y_i$
	Explosives	Romance	Subtitles	Animated	Likes?
e1	T	T	T	T	No
e2	F	T	F	F	Yes
e3	F	F	T	F	No
e4	T	F	T	T	No
e5	T	T	F	T	Yes
e6	T	F	T	F	No



	Explosives	Romance	Subtitles	Animated	Likes?
e7	T	T	F	F	???
e8	F	T	F	F	???

T<sub>1</sub>      T<sub>2</sub>  
 Yes      Yes  
 Yes      Yes

## Decision Tree :-

- Binary Tree
- Internal Nodes are labeled by tests.
- Arcs are labeled by True/False.
- Leaves are labeled by output values.

```

if (subtitles) then
  return No
else
  if (Romance) then
    return Yes
  else
    return No.
  
```

A decision tree can represent any discrete function.

- How to generate trees?
- Which trees should be preferred

**Bias** :- Prefer smaller trees over larger trees.

# Searching for trees:

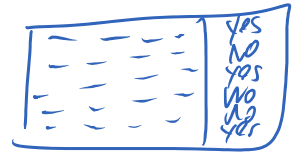
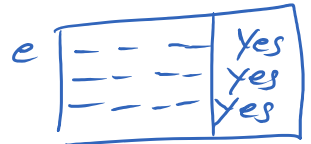
Given examples  $E$   
 if all examples show same output  $Y$   
 return  $Y$

else

- pick a test to split data
- one that divides evenly
  - split gives smallest error.

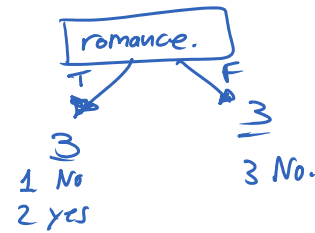
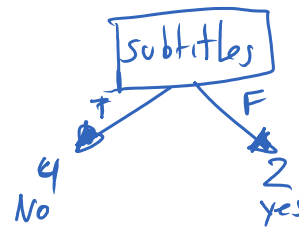
split examples  $e^+$   $e^-$

repeat recursively on  $(e^+)$   
 repeat recursively on  $(e^-)$



## Example:

	Explosives	Romance	Subtitles	Animated	Likes?
e1	T	T	T	T	No
e2	F	T	F	F	Yes
e3	F	F	T	F	No
e4	T	F	T	T	No
e5	T	T	F	T	Yes
e6	T	F	T	F	No



## Research

- How to split data?
- How to compact trees.
- How to handle over-fitting
  - pruning tree when examples are small

