



Data Mining with Weka

More weather

Ian H. Witten



More weather

instances

attributes

	Outlook	Temp	Humidity	Windy	Play
1	Sunny	Hot	High	False	No
2	Sunny	Hot	High	True	No
3	Overcast	Hot	High	False	Yes
4	Rainy	Hot	High	False	Yes
5	Overcast	Cool	Normal	False	Yes
6	Overcast	Cool	Normal	True	No
7	Overcast	Cool	Normal	True	Yes
8	Sunny	Mild	Normal	True	No
9	Sunny	Cool	Normal	False	Yes
10	Rainy	Mild	Normal	False	Yes
11	Sunny	Mild	Normal	True	Yes
12	Overcast	Mild	High	True	Yes
13	Overcast	Hot	Normal	False	Yes
14	Rainy	Mild	High	True	No

Classification problem:
predict the “class” value

More weather

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter: Choose None

Current relation: Relation: weather.symbolic Instances: 14 Attributes: 5

Attributes: All None Invert Pattern

No.	Name
1	<input checked="" type="checkbox"/> outlook
2	<input type="checkbox"/> temperature
3	<input type="checkbox"/> humidity
4	<input type="checkbox"/> windy
5	<input type="checkbox"/> play

Remove

Name: outlook Missing: 0 (0%) Distinct: 3 Type: Nominal Unique: 0 (0%)

No.	Label	Count
1	sunny	5
2	overcast	4
3	rainy	5

Class: play (Nom) Visualize All

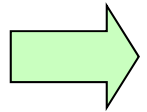
Status: OK Log x 0

More weather

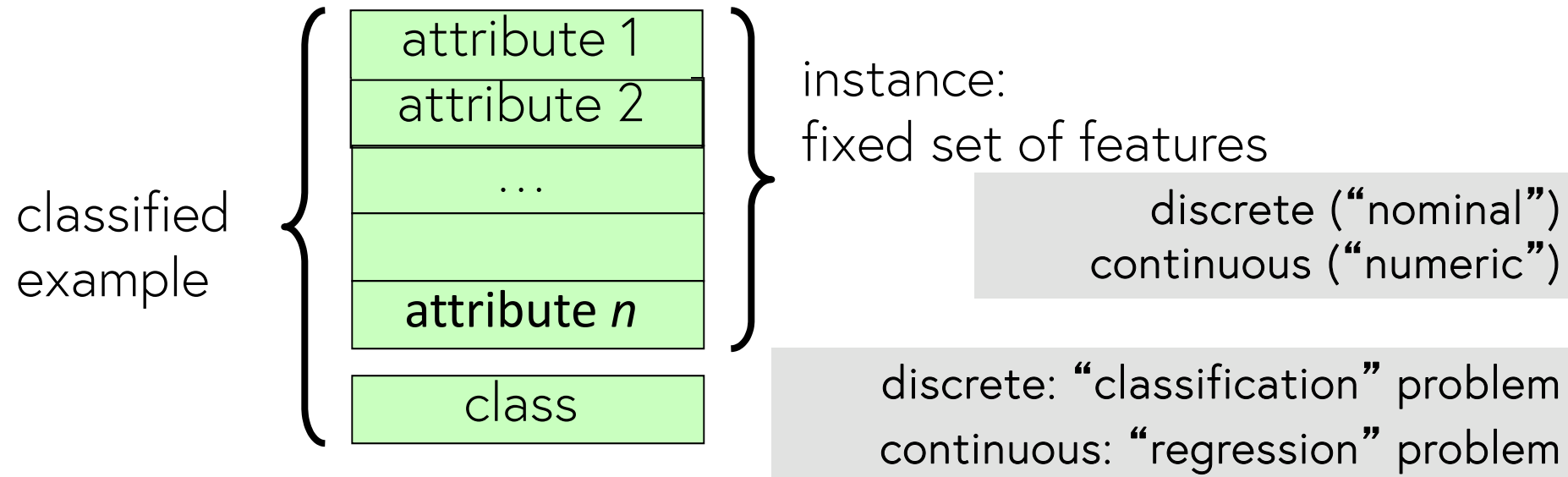
Classification

sometimes called “supervised learning”

Dataset: classified examples



“Model” that classifies new examples



More weather

The screenshot shows the Weka Explorer application window. The 'Preprocess' tab is selected. The 'Open file...' button is highlighted with a callout box containing the text 'open file weather.numeric.arff'. The 'Attributes' list on the left includes 'outlook', 'temperature', 'humidity', 'windy', and 'play', with 'outlook' selected. A callout box labeled 'attributes' points to this list. The 'Current relation' section shows 'Relation: weather.symbolic' and 'Instances: 14'. The 'Attributes: 5' section shows a table with 3 rows: '1 sunny 5', '2 overcast 4', and '3 rainy 5'. A callout box labeled 'attribute values' points to this table. The 'Class: play (Nom)' dropdown is set to 'play (Nom)'. The 'Visualize All' button is visible. The status bar at the bottom shows 'Status OK' and a 'Log' button.

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter
Choose None

Current relation
Relation: weather.symbolic
Instances: 14 Attributes: 5

Attributes
All None Invert Pattern

No.	Name
1	<input checked="" type="checkbox"/> outlook
2	<input type="checkbox"/> temperature
3	<input type="checkbox"/> humidity
4	<input type="checkbox"/> windy
5	<input type="checkbox"/> play

Remove

No.	Label	Count
1	sunny	5
2	overcast	4
3	rainy	5

Class: play (Nom) Visualize All

Status OK Log x 0