Problem 1.

Consider the Wheel Chains Dataset (Table 1) we encountered in the lectures. In this task we will consider any system which had a grade above 2.3 an accepted system, otherwise it is unacceeted.

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	206	1.9	1.4	1.8	2.7	F
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SK	F	160	1.7	1.9	1.6	3.7	F
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SRK	F or R	325	2.2	2.2	4.6	3.2	T
11	SRK	F or R	389	2.0	2.2	3.3	4.3	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	149	1.9	2.5	4.0	3.8	T
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

Table 1 – Wheel Chains Dataset with initial order

1. Consider the last three examples of this dataset only (no cross validation) and classify them using lazy classification with pattern structures (with interval pattern structures for numerical attributes). What is the accuracy of this classification? Why do you think that happend?
2. Now let's consider the same dataset, with a different order (Table 2) below. Try to classify same three examples, what is the accuracy?
3. Assess the performance of the lazy classification with interval pattern structures on the randomly ordered sample (the second one) by applying 5-cross validation.
4. We define a new type of interval pattern structures as follows:

   IF the value of attribute J for object I is $w(I,J)$, the value of the pattern structure will be: $[w(I,J),\infty]$. We will also define the intersection on such intervals as:

   $$[a_1,\infty] \cap [a_2,\infty] = [\min(a_1,a_2),\infty].$$

   Classify the example number 15 using the newly defined pattern structures.
5. We define a new type of interval pattern structures as follows:

   If the value of attribute J for object I is $w(I,J)$, the value of the pattern structure will be: $[w(I,J),\infty]$. We will also define the intersection on such intervals as:

   $$[a_1,\infty] \cap [a_2,\infty] = [\max(a_1,a_2),\infty]$$

   Classify the example number 15 using the newly define pattern structures. Which one of the defined interval pattern structures is most suitable for our dataset? Why?

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

Table 2 – Dataset with different ordering

Solution:

1. Pattern structures used here:

   - For the nominal attributes, we will use the set theoric intersection;
   - For the numerical attributes, we will use the interval pattern structures, where we replace the attribute value $w(I, J)$ where I is the object and J is the attribute, with an interval $[w(I, J), w(I, J)]$

   Let me take the last three examples of the dataset:

№	System	mount	price	con	snow	ice	dur	Accegrade
15	SK	F	99	2.8	2.2	2.5	4.0	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

   The first test example:

   we will loop through all objects in the training dataset and repeat these next steps until we reach a classification:

   - We compute: IntSec $= \delta(g) \cap \delta(g_t) = \langle SK, F, [99, 206], [1.9, 2.8], [1.4, 2.2], [1.8, 2.5], [2.7, 4.0]\rangle$.
   - Now we compute the extension Ext = (IntSec)$'$ (Table 3):

$$\text{Ext} = \langle 1, 6\rangle.$$

   They are both False, then we classify the first object as False.

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	206	1.9	1.4	1.8	2.7	F
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SK	F	160	1.7	1.9	1.6	3.7	F
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SRK	F or R	325	2.2	2.2	4.6	3.2	T
11	SRK	F or R	389	2.0	2.2	3.3	4.3	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	149	1.9	2.5	4.0	3.8	T
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T

Table 3 – Extension for first test example

The second test example

Let's apply the same algorithm to the second test example:

- IntSec $= \langle SK, F, [140, 206], [1.9, 2.6], [1.4, 2.3], [1.8, 3.3], [2.7, 3.4] \rangle$
- Ext $=$ (IntSec)$'$ (Table 4):

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	206	1.9	1.4	1.8	2.7	F
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SK	F	160	1.7	1.9	1.6	3.7	F
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SRK	F or R	325	2.2	2.2	4.6	3.2	T
11	SRK	F or R	389	2.0	2.2	3.3	4.3	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	149	1.9	2.5	4.0	3.8	T
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T

Table 4 – Extension for the second test example

So, we notice that all of them are False, then we classify our example with the same value.

The third test example

Same for the last one in the test set:

- IntSec = $\langle SK, F, [206, 215], [1.9, 2.3], [1.4, 3.8], [1.8, 4.8], [2.3, 2.7] \rangle$
- Ext = (IntSec)′ (Table 5)

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	206	1.9	1.4	1.8	2.7	F
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SK	F	160	1.7	1.9	1.6	3.7	F
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SRK	F or R	325	2.2	2.2	4.6	3.2	T
11	SRK	F or R	389	2.0	2.2	3.3	4.3	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	149	1.9	2.5	4.0	3.8	T
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T

Table 5 – Extension for the third test example

The same as for the second test example. We predict False value of the target variable.

Inferences: our final accuracy is 0%. It became real because of the inital data ordering.

2. Let's do the same for the data with another ordering (Table 2):

The first test example

- IntSec = $\langle SK, F, [99, 149], [1.9, 2.8], [2.2, 2.5], [2.5, 4.0], [3.8, 4.0] \rangle$
- The extension (Table 6):

$$\text{Ext} = \langle 1 \rangle$$

We predict True value for the target variable.

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T

Table 6 – Extension for the first test example for ordered data

The second test example

- IntSec = $\langle SK, F, [140, 149], [1.9, 2.6], [2.3, 2.5], [3.3, 4.0], [3.4, 3.8]\rangle$;
- Extension (Table 7):

$$\text{Ext} = \langle 1\rangle$$

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T

Table 7 – Extension for the second test example for ordered data

Our prediction for the second test example is True.

The third test example

- IntSec = $\langle SK, F, [149, 215], [1.9, 2.3], [2.5, 3.8], [4.0, 4.8], [2.3, 3.8]\rangle$
- Extension (Table 8):

$$\text{Ext} = \langle 1\rangle$$

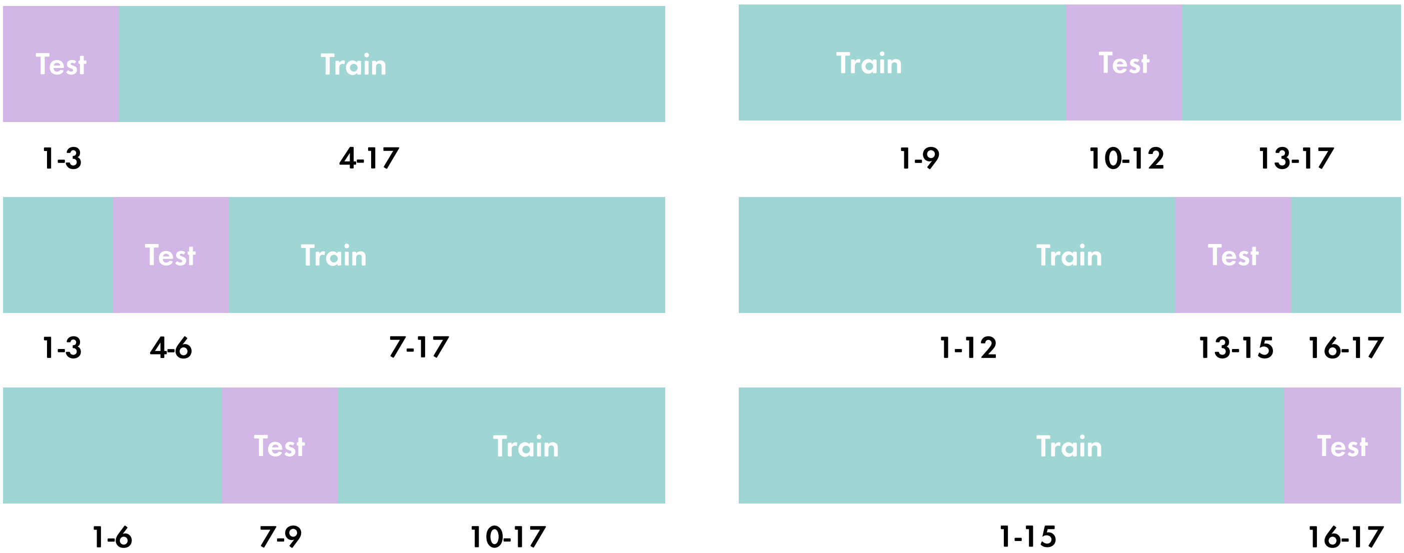
Predicted value for the third test example is True.

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T

Table 8 – Extension for the third test example for ordered data

So we have accuracy is equal to 100%. It very depends on the table order.

3. Let's apply cross-validation method to calculate accuracy of the algorithm. There are 6 folds:



Let's start from the first fold:

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T

And training set:

№	System	mount	price	con	snow	ice	dur	Accegrade
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

For the first example from test set:

- IntSec = $\langle SK, F, [149, 213], [1.7, 1.9], [2.0, 2.5], [2.4, 4.0], [3.4, 3.8]\rangle$
- Ext = $\langle 4\rangle \implies$ False

For the second example from test set:

- IntSec = $\langle *, *, [213, 520], [1.7, 2.1], [0.8, 2.0], [2.4, 3.8], [2.3, 3.4]\rangle$
- Ext = $\langle 4\rangle \implies$ False

For the third example from test set:

- IntSec = $\langle *, *, [213, 389], [1.7, 2.0], [2.0, 2.2], [2.4, 3.3], [3.4, 4.3]\rangle$
- Ext = $\langle 4\rangle \implies$ False

Accuracy 33%.

Second fold:

№	System	mount	price	con	snow	ice	dur	Accegrade
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F

Training set:

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

The first test example:

- IntSec = $\langle SK, F, [149, 213], [1.7, 1.9], [2.0, 2.5], [2.4, 4.0], [3.4, 3.8]\rangle$
- Ext = $\langle 1\rangle \implies$ True

The second test example:

- IntSec = $\langle *, *, [149, 598], [1.6, 1.9], [2.4, 2.5], [4.0, 7], [2.8, 3.8]\rangle$
- Ext = $\langle 1\rangle \implies$ True

The third test example:

- IntSec = $\langle SK, F, [109, 149], [1.9, 2.0], [1.9, 2.5], [2.4, 4.0], [3.7, 3.8]\rangle$

- **Ext** = $\langle 1\rangle \implies$ **True**

Accuracy: 0%.

Third fold:

№	System	mount	price	con	snow	ice	dur	Accegrade
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T

Train set:

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

The first test example:

- **IntSec** = $\langle *, *, [149, 325], [1.9, 2.0], [2.1, 2.5], [3.2, 4.0], [2.8, 3.8]\rangle$
- **Ext** = $\langle 1\rangle \implies$ **True**

The second test example:

- **IntSec** = $\langle *, *, [149, 498], [1.5, 1.9], [2.5, 3.3], [3.5, 4.0], [2.0, 3.8]\rangle$
- **Ext** = $\langle 1\rangle \implies$ **True**

The third test example:

- **IntSec** = $\langle *, *, [149, 396], [1.9, 2.8], [2.1, 2.5], [3.1, 4.0], [2.5, 3.8]\rangle$
- **Ext** = $\langle 1, 12\rangle \implies$ **True**

Accuracy: 67%.

Fourth fold:

№	System	mount	price	con	snow	ice	dur	Accegrade
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T

Train set:

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

The first test example:

- IntSec = $\langle SK, F, [149, 160], [1.7, 1.9], [1.9, 2.5], [1.6, 4.0], [3.7, 3.8]\rangle$
- Ext = $\langle 1\rangle \implies$ True

The second test example:

- IntSec = $\langle *, *, [149, 325], [1.9, 2.2], [2.2, 2.5], [4.0, 4.6], [3.2, 3.8]\rangle$
- Ext = $\langle 1\rangle \implies$ True

The third test example:

- IntSec = $\langle *, F, [149, 298], [1.9, 2.5], [2.3, 2.5], [3.3, 4.0], [2.8, 3.8]\rangle$
- Ext = $\langle 1\rangle \implies$ True

Accuracy: 67%.

Fifth fold:

№	System	mount	price	con	snow	ice	dur	Accegrade
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T

Train set is presented in Table 9:

The first test example:

- IntSec = $\langle SK, F, [149, 206], [1.9, 1.9], [1.4, 2.5], [1.8, 4.0], [2.7, 3.8]\rangle$
- Ext = $\langle 1\rangle \implies$ True

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

Table 9 – CV train set for fifth fold

The second test example:

- IntSec = $\langle *, *, [149, 684], [1.7, 1.9], [2.5, 3.3], [4.0, 4.4], [2.2, 3.8] \rangle$
- Ext = $\langle 1 \rangle \implies$ True

The third test example:

- IntSec = $\langle SK, F, [99, 149], [1.9, 2.8], [2.2, 2.5], [2.5, 4.0], [3.8, 4.0] \rangle$
- Ext = $\langle 1 \rangle \implies$ True

Accuracy: 67%.

Last fold:

№	System	mount	price	con	snow	ice	dur	Accegrade
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

Train set is presented in Table 10

The first test example:

- IntSec = $\langle SK, F, [140, 149], [1.9, 2.6], [2.3, 2.5], [3.3, 4.0], [3.4, 3.8] \rangle$
- Ext = $\langle 1 \rangle \implies$ True

The second test example:

- IntSec = $\langle SK, F, [149, 215], [1.9, 2.3], [2.5, 3.8], [4.0, 4.8], [2.3, 3.8] \rangle$
- Ext = $\langle 1 \rangle \implies$ True

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
15	SK	F	99	2.8	2.2	2.5	4.0	T

Table 10 – CV train set for sixth fold

Accuracy: 100%.

Accuracy through cross-validation: 55.67%.

4. New definition of intersection on interval patterns:

$$[a_1, \infty] \cap [a_2, \infty] = [\min(a_1, a_2), \infty].$$

Let's classify the fifteenth object. Train set

№	System	mount	price	con	snow	ice	dur	Accegrade
1	SK	F	149	1.9	2.5	4.0	3.8	T
2	SRK	F or R	520	2.1	0.8	3.8	2.3	F
3	SRK	F or R	389	2.0	2.2	3.3	4.3	T
4	SK	F	213	1.7	2.0	2.4	3.4	F
5	SMS	F or R	598	1.6	2.4	7	2.8	F
6	SK	F	109	2.0	1.9	2.4	3.7	F
7	SRK	F or R	325	2.0	2.1	3.2	2.8	F
8	SMS	F or R	498	1.5	3.3	3.5	2.0	T
9	SRK	F or R	396	2.8	2.1	3.1	2.5	T
10	SK	F	160	1.7	1.9	1.6	3.7	F
11	SRK	F or R	325	2.2	2.2	4.6	3.2	T
12	SRK	F	298	2.5	2.3	3.3	2.8	T
13	SK	F	206	1.9	1.4	1.8	2.7	F
14	SMS	F or R	684	1.7	3.3	4.4	2.2	T
16	SK	F	140	2.6	2.3	3.3	3.4	T
17	SK	F	215	2.3	3.8	4.8	2.3	T

№	System	mount	price	con	snow	ice	dur	Accegrade
15	SK	F	99	2.8	2.2	2.5	4.0	T

- **IntSec** = $\langle SK, F, [99, \infty], [1.9, \infty], [2.2, \infty], [2.5, \infty], [3.8, \infty]\rangle$
- **Ext** = $\langle 1 \rangle \Longrightarrow$ **True**

5. **Another definition of intersection on interval patterns:**

$$[a_1, \infty] \cap [a_2, \infty] = [\max(a_1, a_2), \infty].$$

- **IntSec** = $\langle SK, F, [149, \infty], [2.8, \infty], [2.5, \infty], [4.0, \infty], [4.0, \infty]\rangle$
- **Ext** = $\langle 1 \rangle \Longrightarrow$ **True**

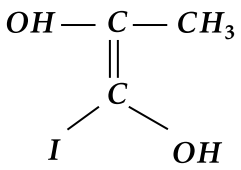
The definition from the 4th subtask is more suitable for our dataset. In general, if the test example will have the maximum element, then we will obtain undefined extension.

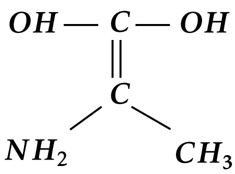
■

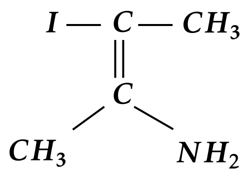
Problem 2.

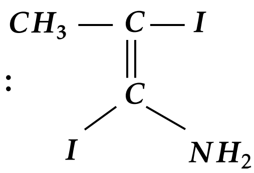
Given the molecular graphs of 4 example substances belonging to the positive class and 3 example substances belonging to the negative class. Using the JSM method, construct pattern concept lattices for the positive (objects $G_1, \dots, G_4$) and negative (objects $G_5, \dots, G_7$) contexts, find their respective minimum positive and negative hypotheses and classify the test examples (objects $G_8, \dots, G_{11}$), accordingly.

Positive examples

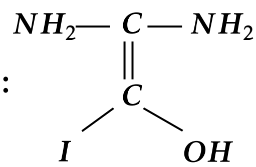
G_1 :


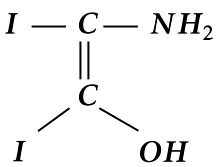
G_2 :


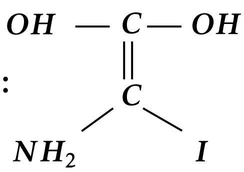
G_3 :


G_4 :


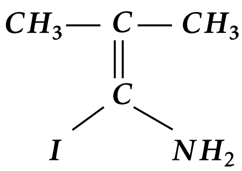
Negative examples

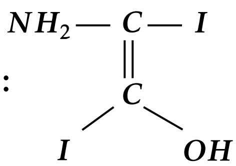
G_5 :


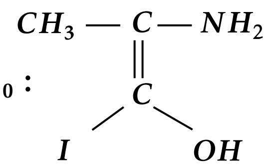
G_6 :


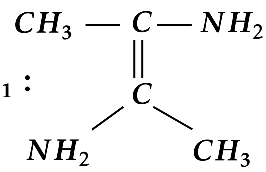
G_7 :


Test examples

G_8 :


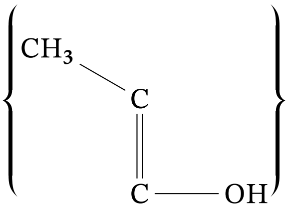
G_9 :


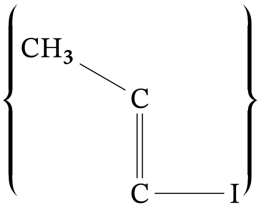
G_{10} :


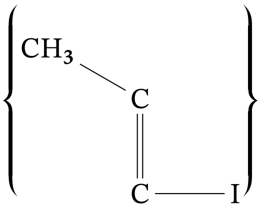
G_{11} :


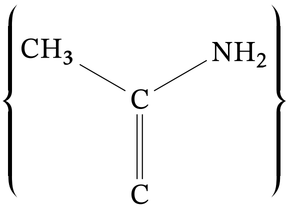
Solution:

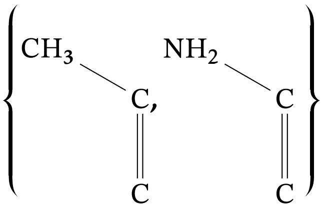
Positive

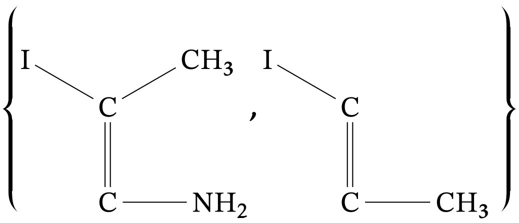
$\{G_1, G_2\}^\diamond = G_1^\diamond \cap G_2^\diamond =$ 
$, \quad \{G_1, G_2\}^{\diamond\diamond} = \{G_1, G_2\}$

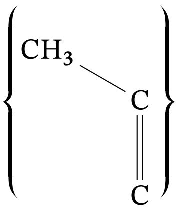
$\{G_1, G_3\}^\diamond =$ 
$, \quad \{G_1, G_3\}^{\diamond\diamond} = \{G_1, G_3, G_4\}$

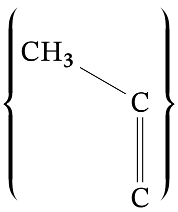
$\{G_1, G_4\}^\diamond =$ 
$, \quad \{G_1, G_4\}^{\diamond\diamond} = \{G_1, G_3, G_4\}$

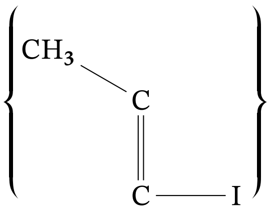
$\{G_2, G_3\} =$ 
$, \quad \{G_2, G_3\}^{\diamond\diamond} = \{G_2, G_3\}$

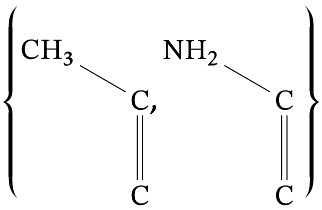
$\{G_2, G_4\} =$ 
$, \quad \{G_2, G_4\}^{\diamond\diamond} = \{G_2, G_3, G_4\}$

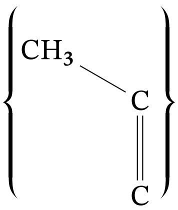
$\{G_3, G_4\} =$ 
$, \quad \{G_3, G_4\}^{\diamond\diamond} = \{G_3, G_4\}$

$\{G_1, G_2, G_3\} =$ 
$, \quad \{G_1, G_2, G_3\}^{\diamond\diamond} = \{G_1, G_2, G_3, G_4\}$

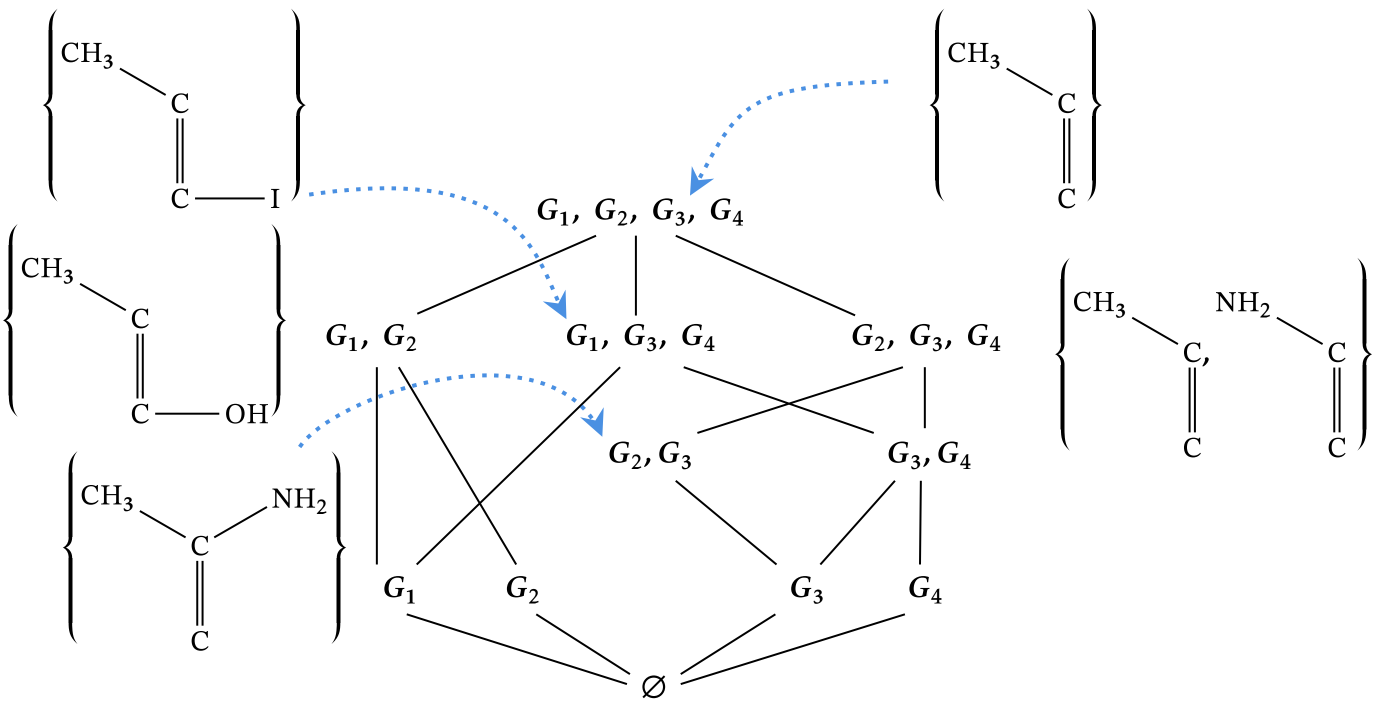
$\{G_1, G_2, G_4\} =$ 
$, \quad \{G_1, G_2, G_4\}^{\diamond\diamond} = \{G_1, G_2, G_3, G_4\}$

$\{G_1, G_3, G_4\} =$  $, \quad \{G_1, G_3, G_4\}^{\diamond\diamond} = \{G_1, G_3, G_4\}$

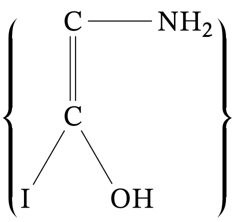
$\{G_2, G_3, G_4\} =$  $, \quad \{G_2, G_3, G_4\}^{\diamond\diamond} = \{G_2, G_3, G_4\}$

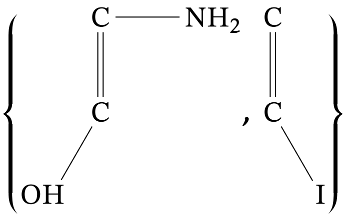
$\{G_1, G_2, G_3, G_4\} =$  $, \quad \{G_1, G_2, G_3, G_4\}^{\diamond\diamond} = \{G_1, G_2, G_3, G_4\}$

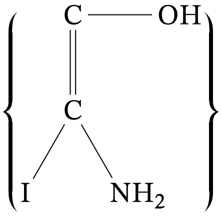
Pattern concept lattice:

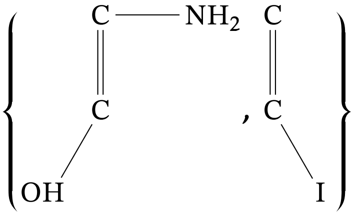


Negative

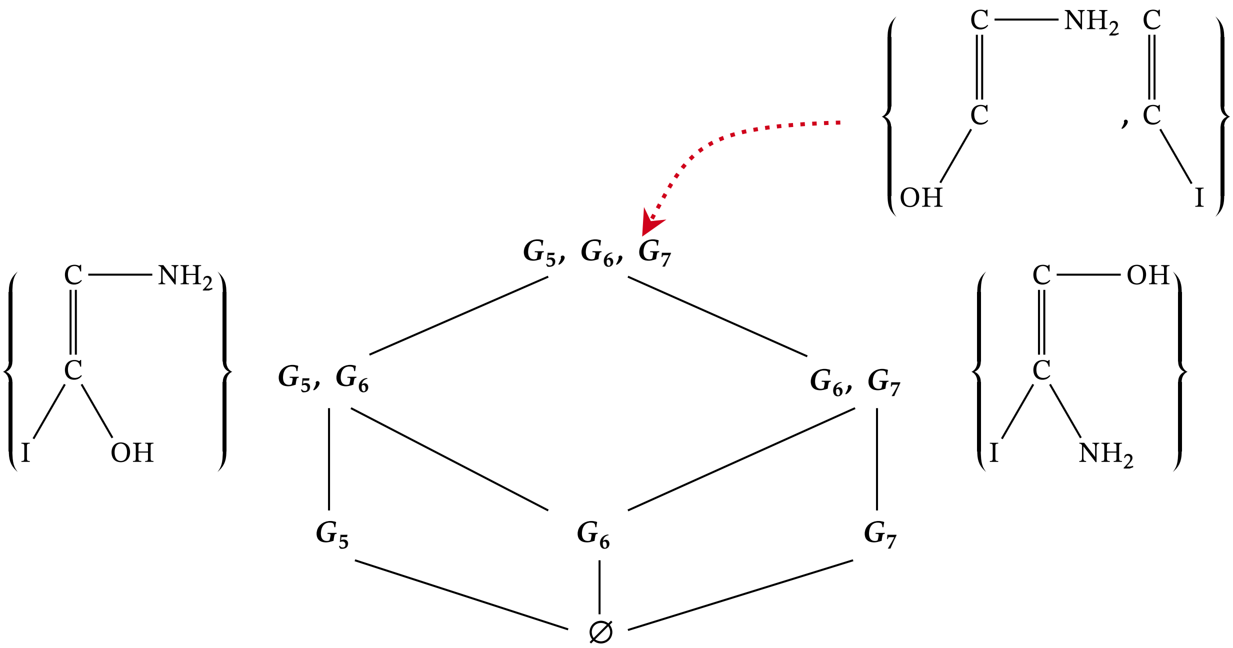
$\{G_5, G_6\}^{\diamond} = G_5^{\diamond} \cap G_6^{\diamond} =$  $, \quad \{G_5, G_6\}^{\diamond\diamond} = \{G_5, G_6\}$

$\{G_5, G_7\}^\diamond = G_5^\diamond \cap G_7^\diamond =$  , $\{G_5, G_7\}^{\diamond\diamond} = \{G_5, G_6, G_7\}$

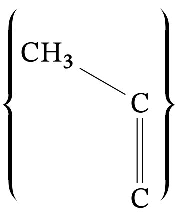
$\{G_6, G_7\}^\diamond = G_6^\diamond \cap G_7^\diamond =$  , $\{G_6, G_7\}^{\diamond\diamond} = \{G_6, G_7\}$

$\{G_5, G_6, G_7\}^\diamond = G_5^\diamond \cap G_6^\diamond \cap G_7^\diamond =$  , $\{G_5, G_6, G_7\}^{\diamond\diamond} = \{G_5, G_6, G_7\}$

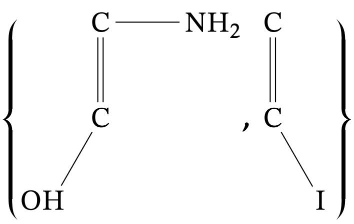
Pattern concept lattice:

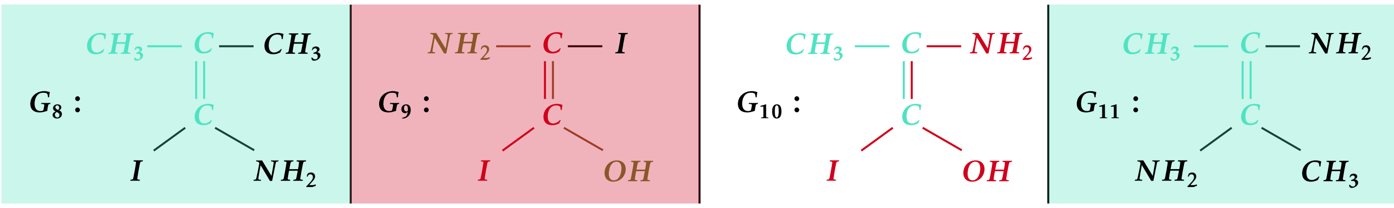


Minimal positive hypothesis:

$H^+ :$ 

Minimal negative hypothesis:

$H^- :$ 



- Object G_8 is classified positively, because:

$$H^+ \sqsubseteq G_8^\diamond,\ H^- \not\sqsubseteq G_8^\diamond$$

- Object G_9 is classified negatively, because:

$$H^- \sqsubseteq G_9^\diamond,\ H^+ \not\sqsubseteq G_9^\diamond$$

- Object G_{10} is classified contradictory, because:

$$H^-, H^+ \sqsubseteq G_{10}^\diamond$$

- Object G_{11} is classified positively, because:

$$H^+ \sqsubseteq G_{11}^\diamond,\ H^- \not\sqsubseteq G_{11}^\diamond$$

■

G_1 :