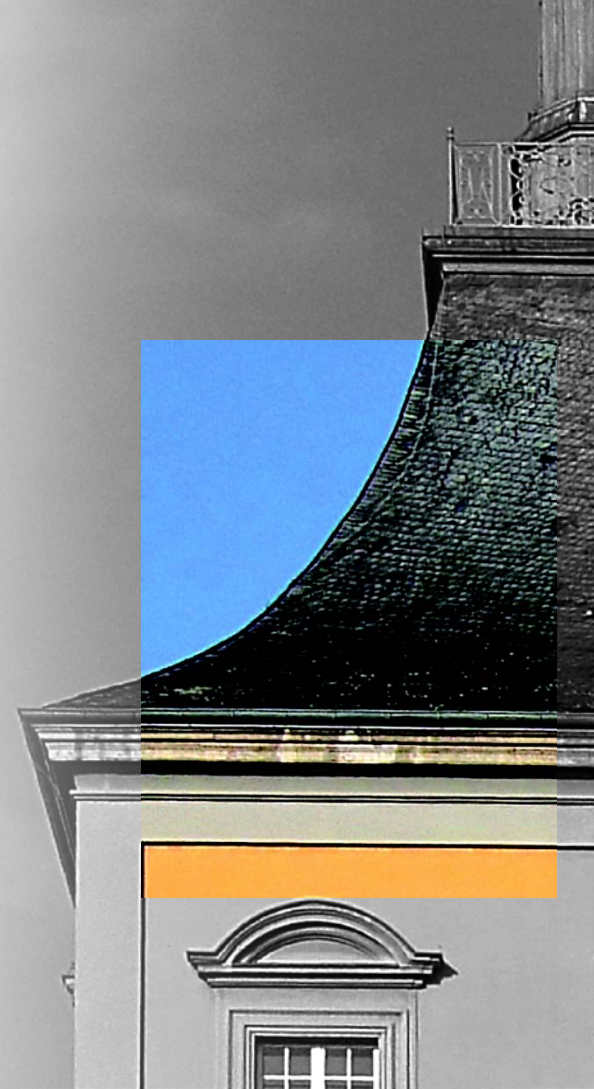


ILAS EXERCISE 2

GROUP 20

MARTIN SCHMITZ
PAUL BRYAN ALEXANDER HINZEN



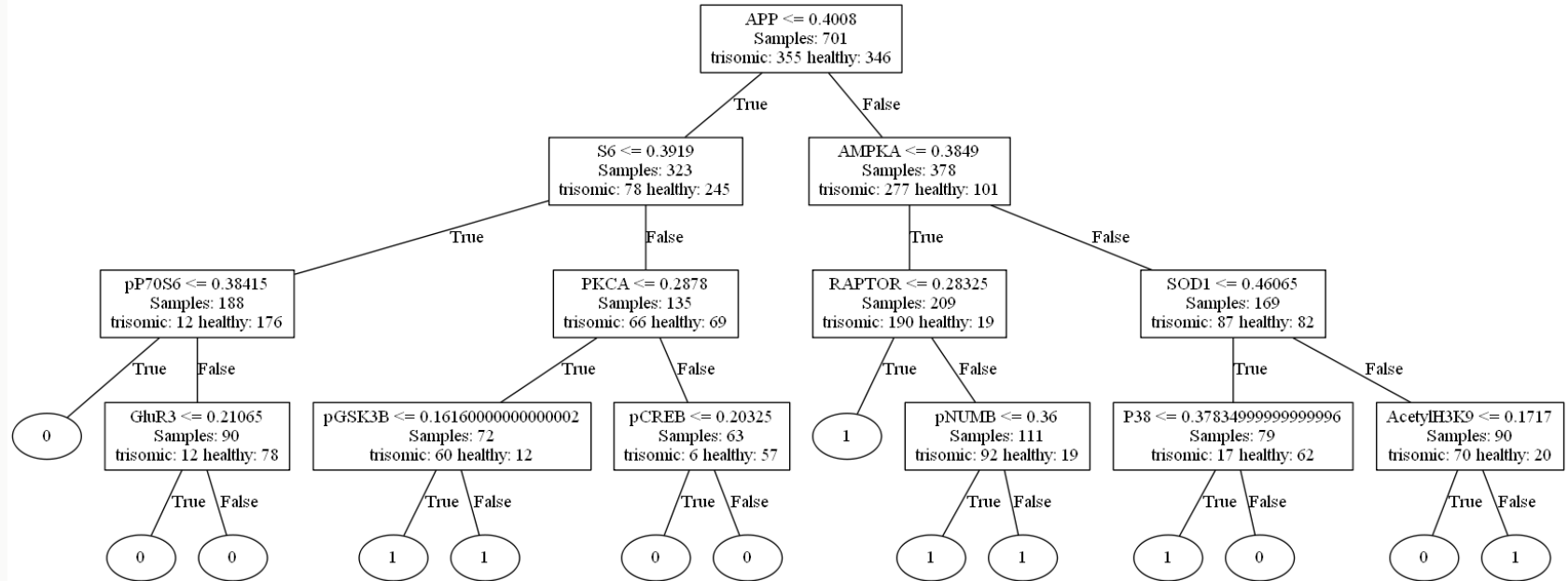
USED TECHNOLOGIES

- Python
- Simple Tree structure
- Pandas DataFrame
- Used Libraries:
 - NumPy
 - Pandas
 - PyDot
 - Pickle

```
class Node(object):
    def __init__(self, used_atts = [], examples = None,
                  leaf = False, label = None):
        self.used_atts = used_atts
        self.examples = examples
        self.left = None
        self.right = None
        self.split_attr = None
        self.split_value = None
        self.label = label
        self.leaf = leaf

    def becomes_leaf(self, c):
        self.leaf = True
        self.label = c
```

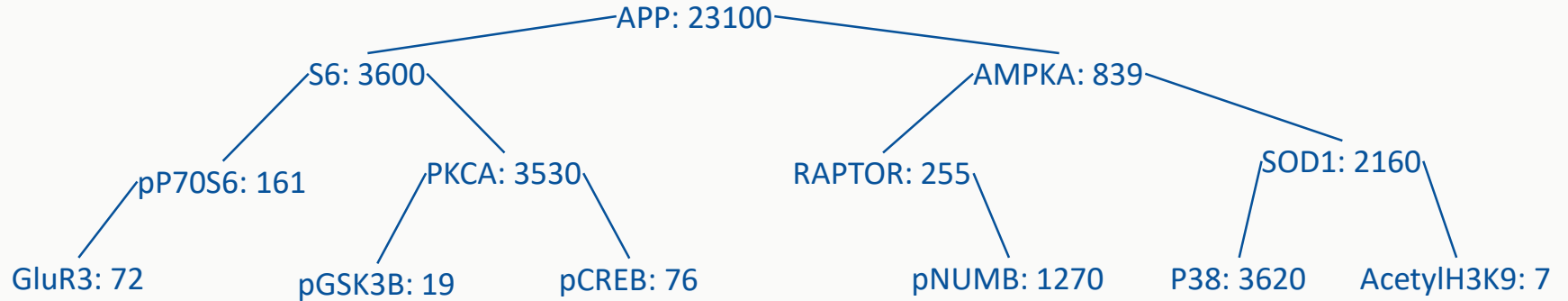
USED TECHNOLOGIES



– Accuracy: 0.853 (Depth 3)

Accuracy: 0.922 (Depth 7)

GOOGLE SCHOLAR SEARCH

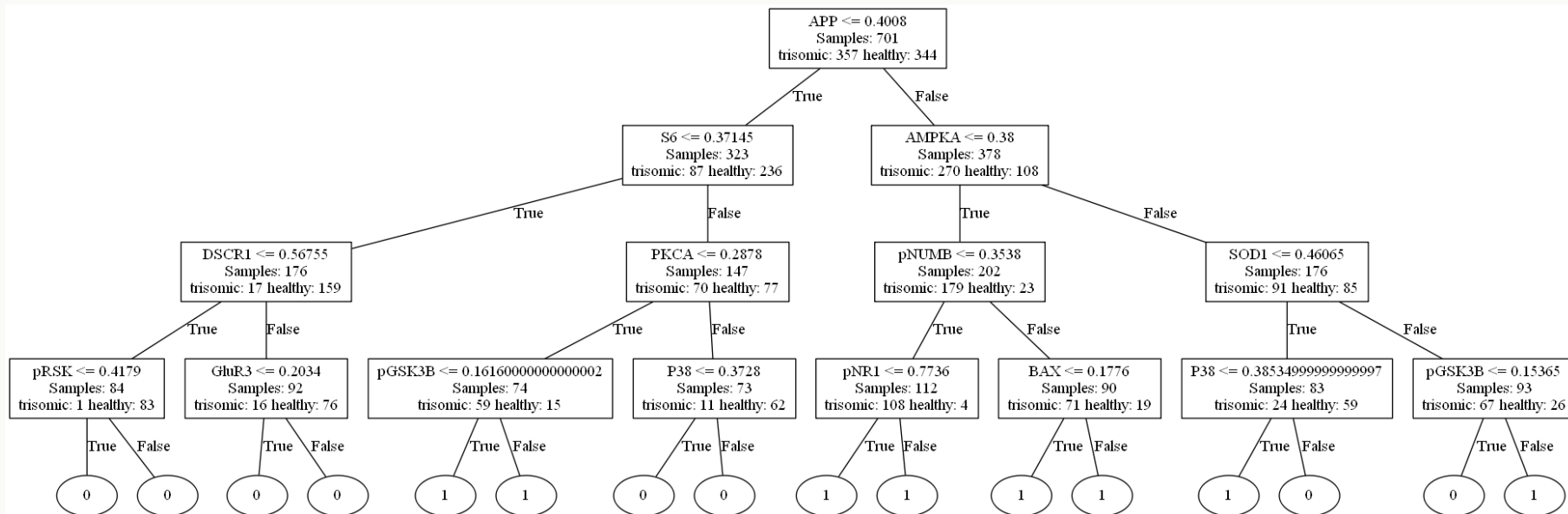




– Effect of label noise:

- 1. “Flip” each class label of the training data with a probability of
 - (a) 0.1, and
 - (b) 0.25, respectively.
 - 2. Construct a decision tree using the training data with “noisy” labels.
 - 3. Measure the accuracy of the decision tree on the test data. (Please note that you should not flip the labels of the test data!)
-

GROUP DEPENDENT TASK



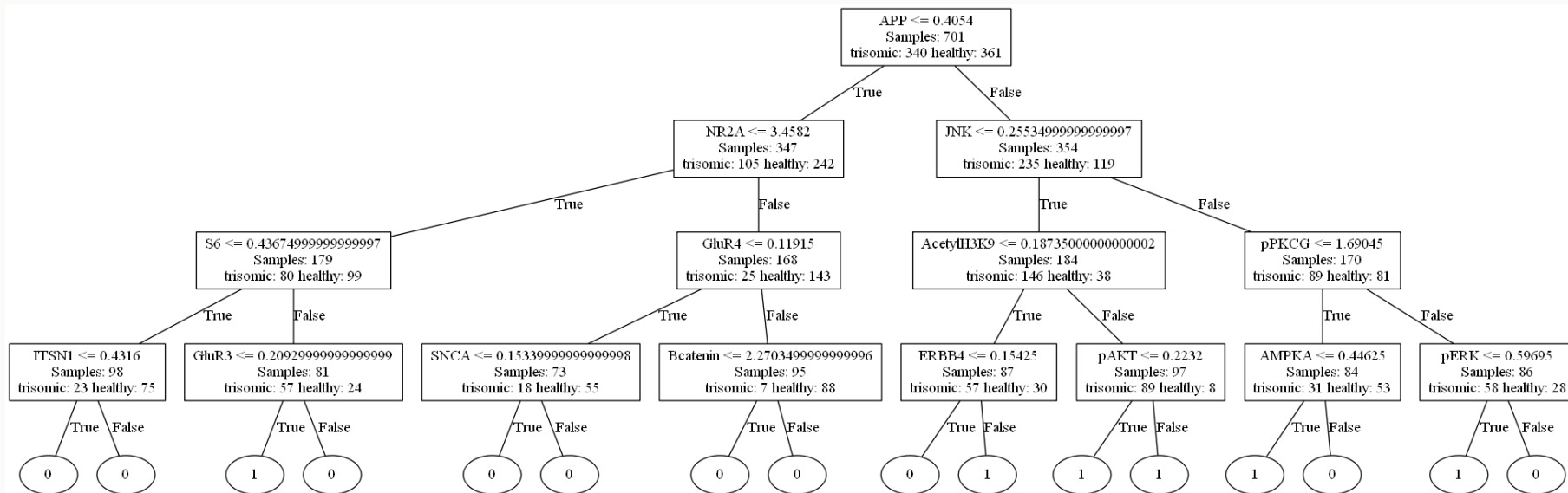
– Labels Flipped: 61 / 701

Percentage: 0.087

– Accuracy: 0.829 (Depth 3)

Accuracy: 0.87 (Depth 7)

GROUP DEPENDENT TASK



– Labels Flipped: 194 / 701

Percentage: 0.2767475035663338

– Accuracy: :0.795 (Depth 3)

Accuracy: 0.777 (Depth 7)

Thank you for your kind attention.

Martin Schmitz

Paul Bryan Alexander Hinzen

Group 20