**CST8390 - Lab 3**

**Classification by Decision Trees**

**Student Name: Hasibullah Yosufi**

**Student Number: 041012318**

2.

A: number of instances: 768

B: number of attributes: 9

C:

* Number of times pregnant
* Plasma glucose concentration (2 hours in an oral glucose tolerance test)
* Diastolic blood pressure
* Triceps skin fold thickness
* 2-hour serum insulin
* Body mas index
* Diabetes pedigree
* Age
* class

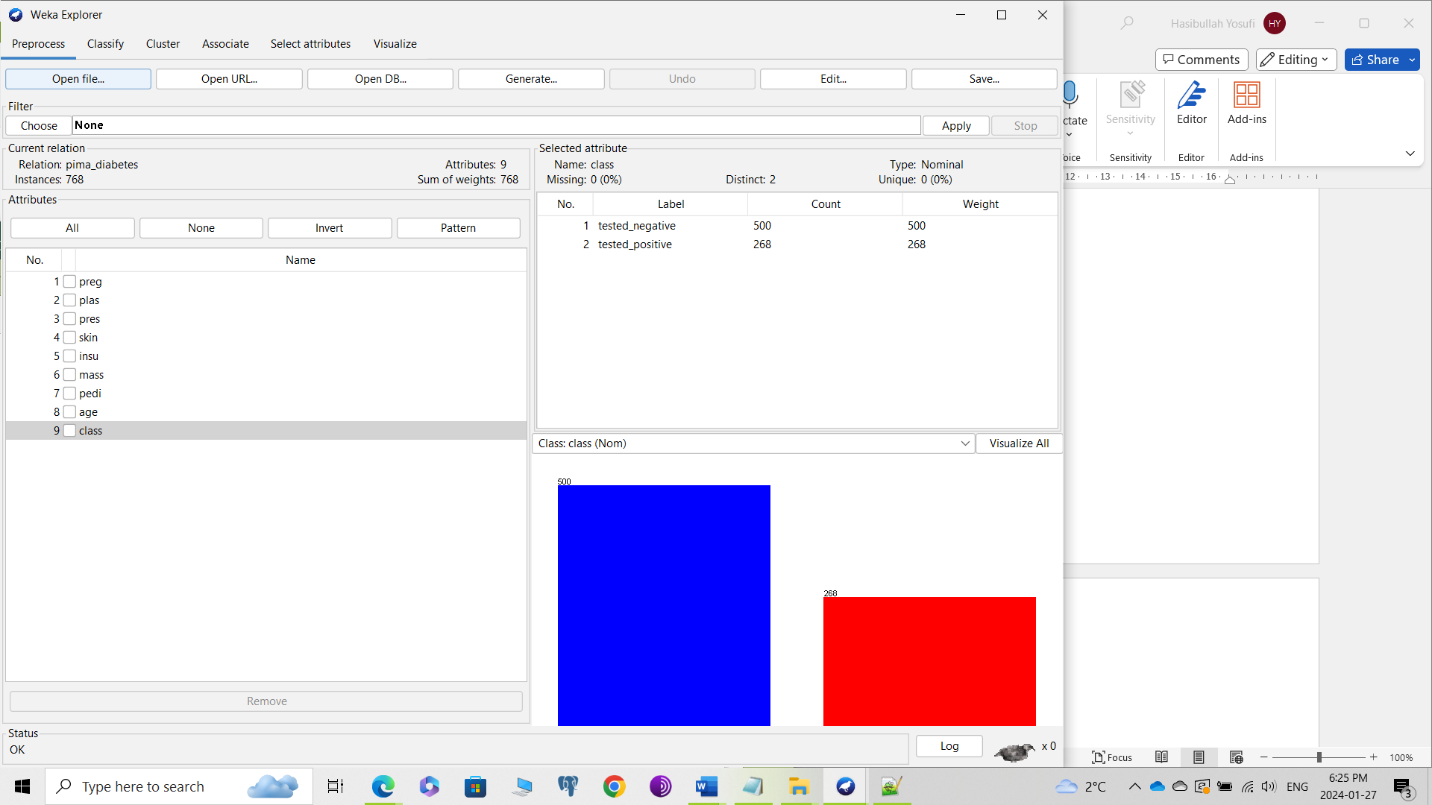
)

D:

tested\_negative (0) , tested\_positive (1

E:

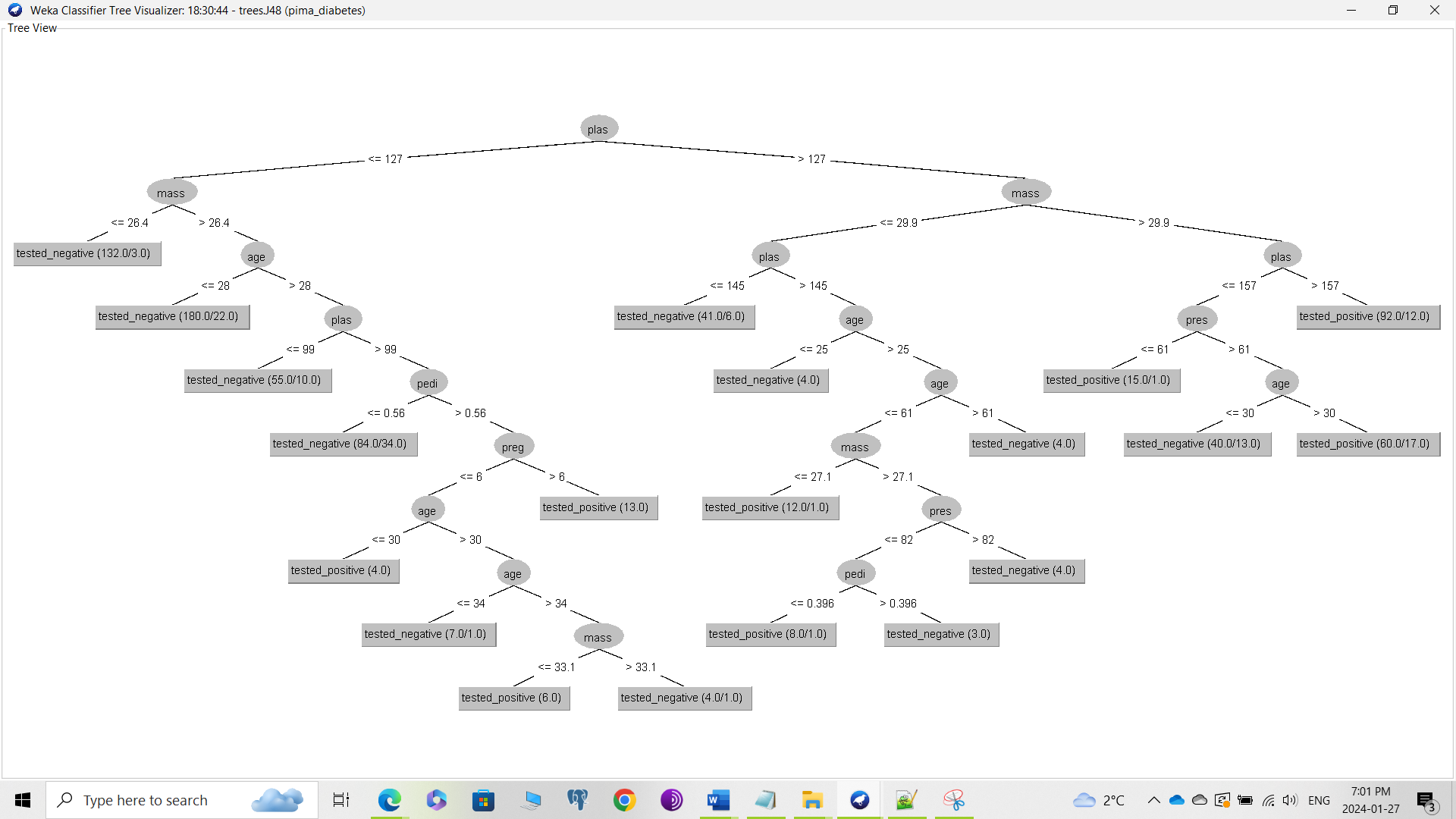
* tested\_negative (0): 500
* tested\_positive (1): 268



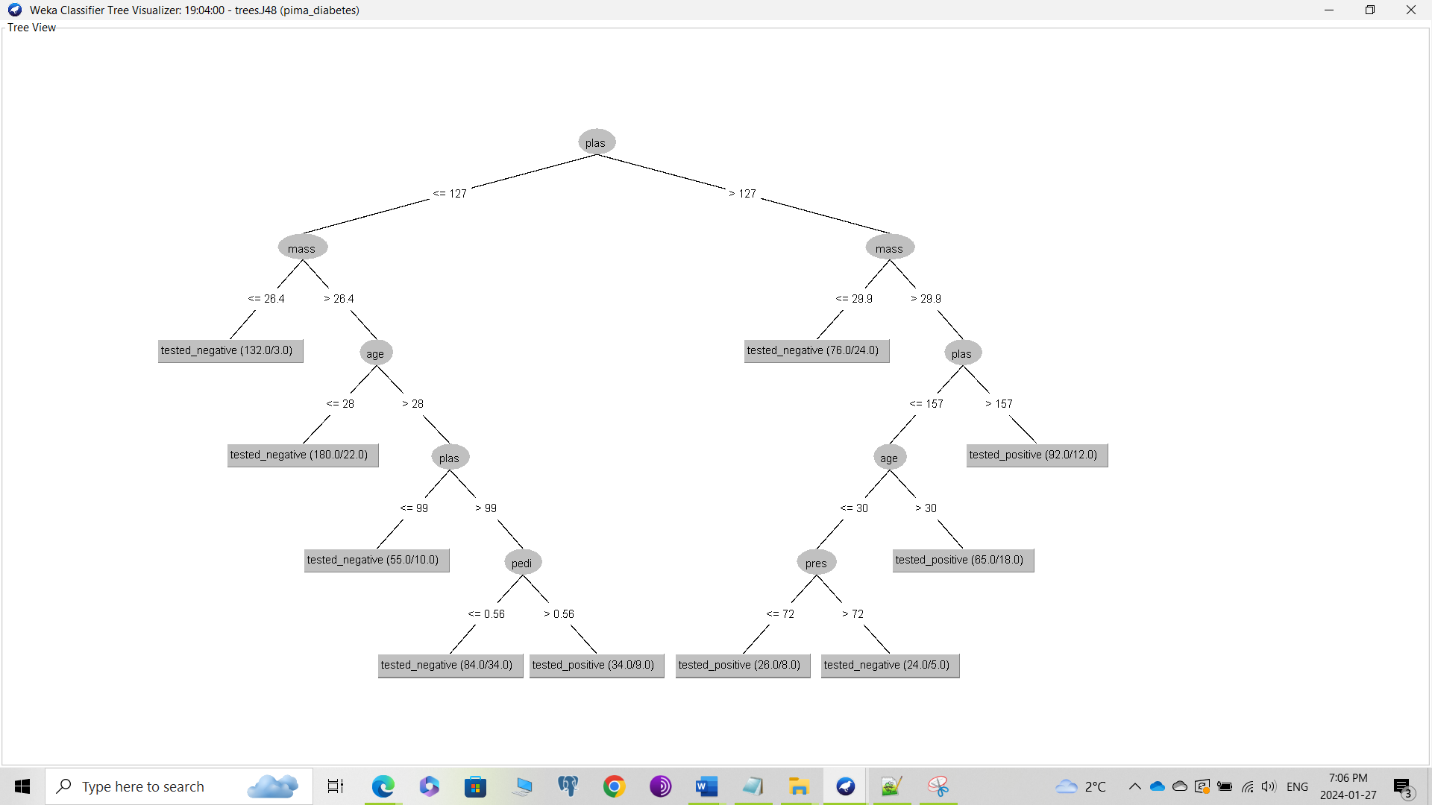
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Question** | **Confusion Matrix** | **Number of leaves** | **Size of the tree** | **Correctly classified instances** |
| 4 |  | 20 | 39 | 577 |
| 6 |  | 10 | 19 | 561 |
| 8 |  | 12 | 23 | 577 |
| 9 |  | 10 | 19 | 569 |

5.

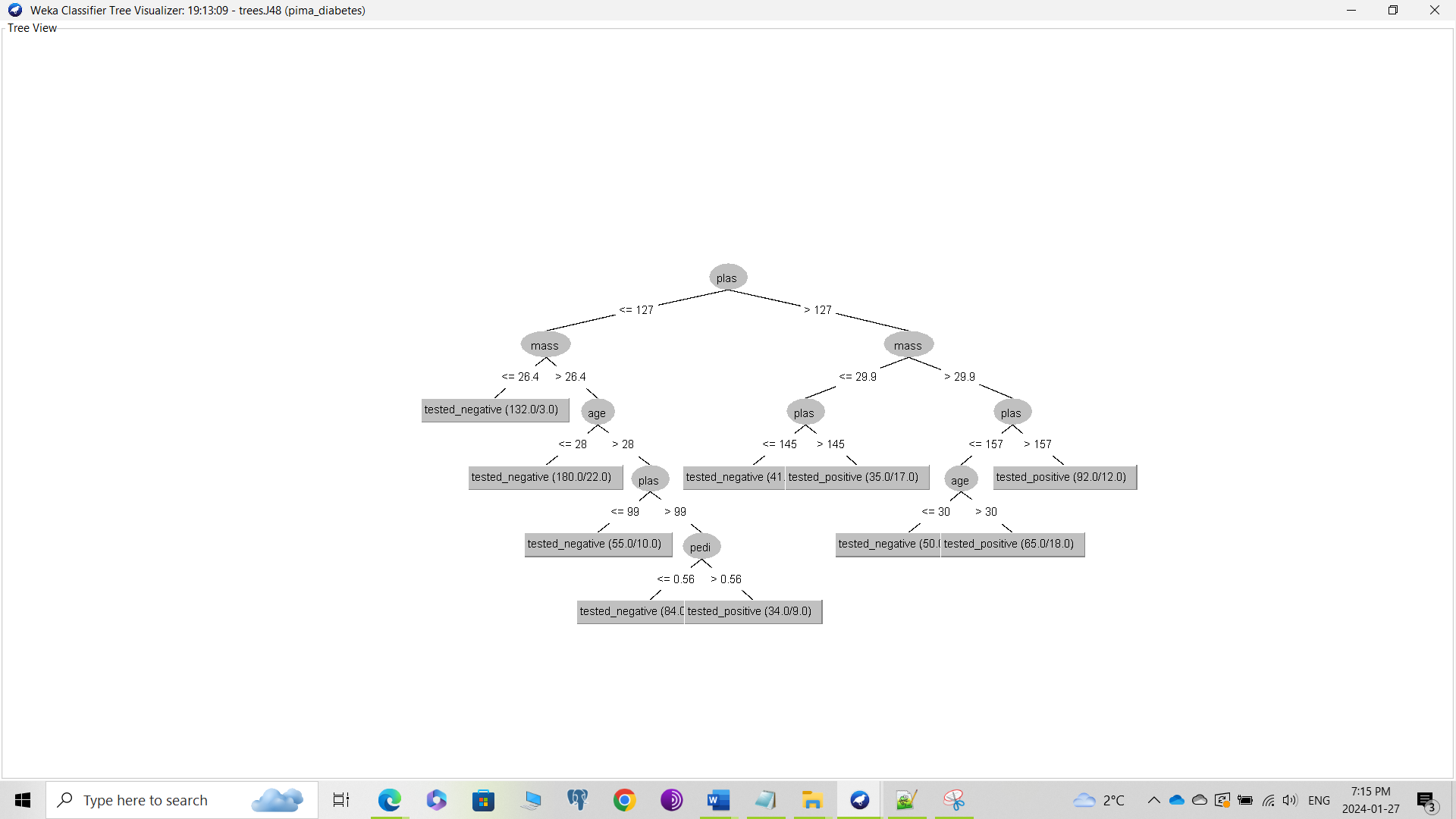
Decision tree:



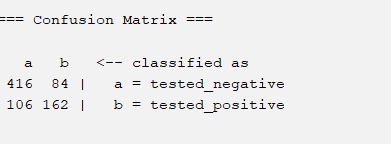
7.



10:



11.



12.

1. Details of random forest: Bagging with 100 iterations
2. Time taken to build the model: 1.01s
3. Correctly classified instances: 578

**Screenshots with Question Number**