**Results**

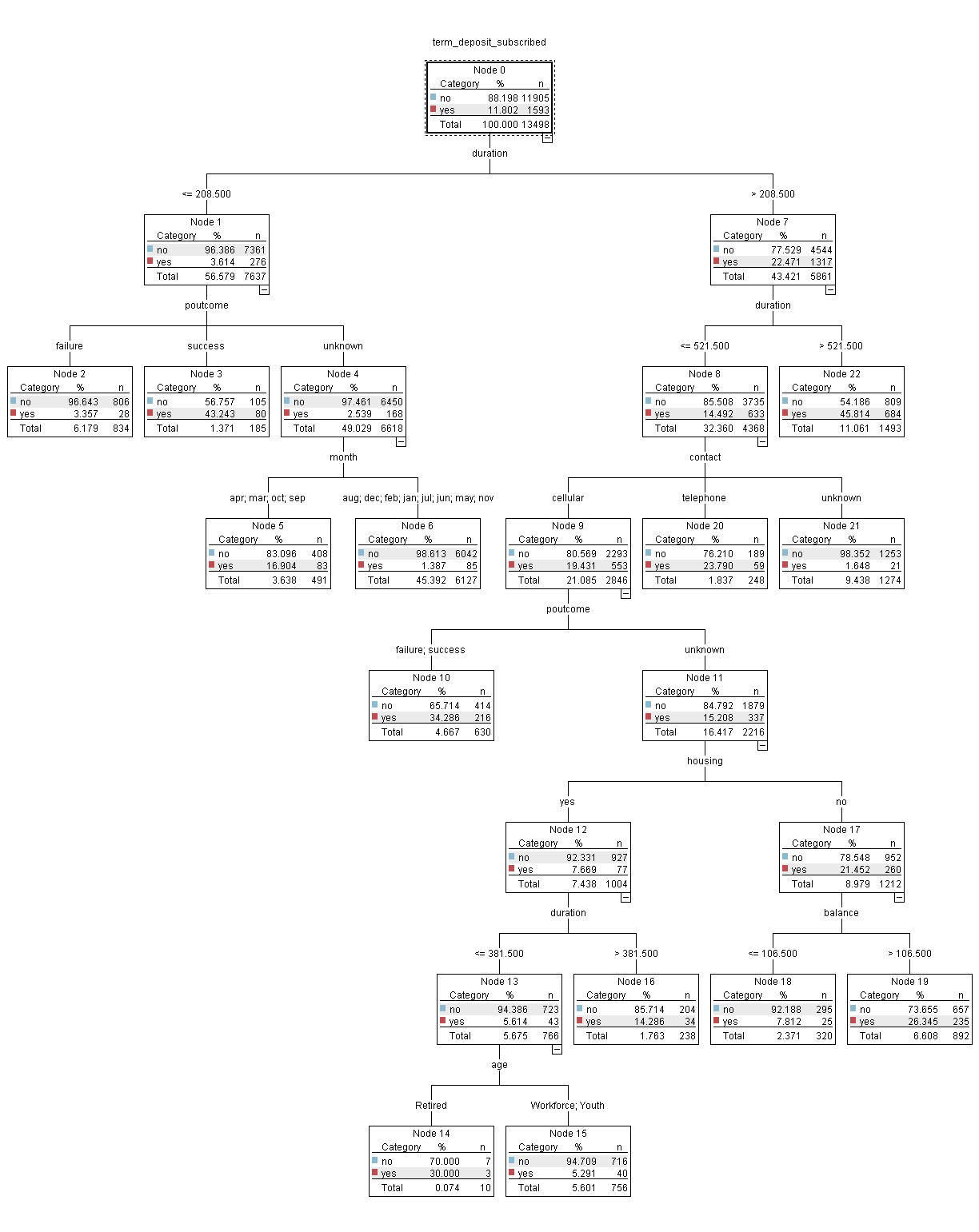
The best model is C5.0 with 30% pruning severity, 500 minimum nodes in child branch, with global pruning, with cost ratio of 9:1, with overall accuracy of 74.5%, with 73% accuracy and lift 1.11 for class ‘no’ and with 89% accuracy and lift 2.6 for class ‘yes’. We wanted to correctly identify ‘yes’ labels so we compromised the overall accuracy and class ’no’ accuracy. The accuracies and lift were consistent throughout all the dataset so we can rely on the model. We can see results in Fig *NUMBER.*

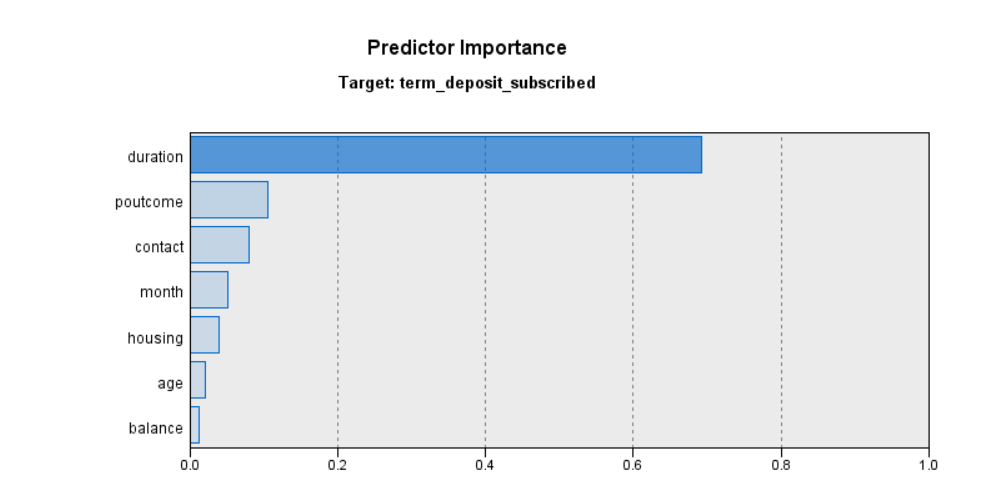
Many predictors were given as an input but we can see that the 3 most important of them are duration, poutcome and contact in Fig *NUMBER*. We can also see most important rules in the table below[For all rules Table *NUMBER*].

|  |  |  |
| --- | --- | --- |
| **#Rule** | **Rule statement** | **Classified as** |
| 1 | duration <= 208.5 and poutcome = ‘failure’ | no |
| 2 | duration <= 208.5 and poutcome = ‘success’ | yes |
| 3 | duration <= 208.5 and duration > 521.5 | yes |
| 4 | duration <= 208.5 and duration <= 521.5 and contact = ‘telephone’ | yes |
| 5 | duration <= 208.5 and duration <= 521.5 and contact = ‘unknown’ | no |
| 6 | duration <= 208.5 and duration <= 521.5 and contact = ‘cellular’ and poutcome = [‘failure’, ‘success’] | yes |

**Table *NUMBER***

First rule is obvious as duration and poutcome were the most important predictors. And also, if duration of the call is shorter, it seems that client is not interested in the campaign details and s/he is not looking forward to subscribing the term deposit. Same goes with the result of previous campaign(poutcome), if a client didn’t invest in the previous campaign s/he is more likely not to subscribe the term deposit. So that’s why those observations are labelled as ‘no’. Second rule is the exact opposite of the former one and we can see that it is classified as ‘yes’. Third rule states the same that higher the duration higher the interest of the client in the campaign and higher the chances of the client subscribing the term deposit. Now the third most important predictor comes into the picture which is contact. Forth rule indicates that if communication medium is telephone and duration is high then those observation can be classified as ‘yes’. The clients are using telephone that suggests that either clients are retired or housewives or someone who stays at home. As these clients have money and easy to convince, so they are labelled as ‘yes’. Fifth rule same as forth rule, different in the terms of only communication medium which is ‘unknown’. Unknown can mean many things that client was not contacted or contacted through mediums like word of mouth by bank employee, other platforms like bank’s mobile application or emails. They are labelled as ‘no’. Sixth rule is the combination of all the above rules except the communication medium is ‘cellular’ and poutcome is either ‘failure’ or ‘success’ and they are labelled as ‘yes’. These rules show that duration, poutcome and the contact is the most associated with the outcome whether client subscribes a term deposit or not.





|  |  |  |
| --- | --- | --- |
| **#Rule** | **Rule statement** | **Classified as** |
| 1 | duration <= 208.5 and poutcome = ‘failure’ | no |
| 2 | duration <= 208.5 and poutcome = ‘success’ | yes |
| 3 | duration <= 208.5 and poutcome = ‘unknown’ and month = [‘apr’, ‘march’, ‘oct’, ‘sep’] | yes |
| 4 | duration <= 208.5 and poutcome = ‘unknown’ and month != [‘apr’, ‘march’, ‘oct’, ‘sep’] | no |
| 5 | duration <= 208.5 and duration > 521.5 | yes |
| 6 | duration <= 208.5 and duration <= 521.5 and contact = ‘telephone’ | yes |
| 7 | duration <= 208.5 and duration <= 521.5 and contact = ‘unknown’ | no |
| 8 | duration <= 208.5 and duration <= 521.5 and contact = ‘cellular’ and poutcome = [‘failure’, ‘success’] | yes |
| 9 | duration <= 208.5 and duration <= 521.5 and contact = ‘cellular’ and poutcome = ‘unknown’ and housing = ‘no’ and balance <= 106.5 | no |
| 10 | duration <= 208.5 and duration <= 521.5 and contact = ‘cellular’ and poutcome = ‘unknown’ and housing = ‘no’ and balance > 106.5 | yes |
| 11 | duration <= 208.5 and duration <= 521.5 and contact = ‘cellular’ and poutcome = ‘unknown’ and housing = ‘yes’ and duration > 381.5 | yes |
| 12 | duration <= 208.5 and duration <= 521.5 and contact = ‘cellular’ and poutcome = ‘unknown’ and housing = ‘yes’ and duration <= 381.5 and age = ‘Retired’ | yes |
| 13 | duration <= 208.5 and duration <= 521.5 and contact = ‘cellular’ and poutcome = ‘unknown’ and housing = ‘yes’ and duration <= 381.5 and age = [‘Workforce’, ‘Youth’] | no |

