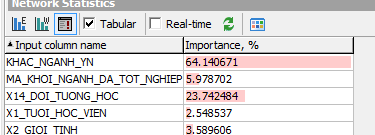
1. Tiền xử lý
2. OutPut: Pass qua kỳ 3(1) + chưa pass qua kỳ 3 (0) (Lấy OutPut dựa vào kỳ đóng học phí gần nhất và trạng thái học viên)

Kỳ(0,1,2,3) and Trạng thái (11) : 0

Kỳ(4,5,6,7) and (8,9) : 1

1. Input
2. Tuổi
3. Giới tính
4. Tôn giáo (Search Excel và sửa giá trị nhiễu)
5. Đối tượng học: Đã học cái gì đó rồi!
6. Bằng tốt nghiệp cấp 3: Quá nhiều nhiễu , ko biết chỉnh kiểu gì =>> Bỏ
7. Nghành học: Bỏ( không ý nghĩa với Profile mới, chưa học gì)
8. Nơi sinh: Thử cho vào xem thế nào :D
9. Các loại điểm đầu vào: Đã bỏ đi ( Mình nghĩ cái này nên cho vào để học viên cảm thấy quý giá khi được học, có lẽ do mô hình của họ thời điểm này chưa cần nó) =>>Bỏ
10. Khác nghành YN (Có những cái chưa biết thì sao? Gán vào: chưa biết)
11. Mã khối nghành đã tốt nghiệp
12. Đã có QĐ: bỏ
13. Đã TN: bỏ( toàn N)
14. Ngày tháng năm sinh: Hiện tại bỏ(giống tuổi)? Liệu có tính theo tháng được ko? Tính theo ngày? (Liên quan đến **Nhân tướng học**)
15. Phân tích tương quan (Phần trăm input quan trọng ảnh hưởng đến đầu ra. Không biết bọn nó tính thế nào nhỉ?)



1. Kết quả chạy thử trên Alyuda



1. Tiếp theo nên làm gì để tăng độ chính xác của kết quả dự báo nên 5%?
2. Thêm input: Sửa một số input thiếu, thêm điểm:( Khai phá bảng điểm) ,chạy xem thế nào
3. Chạy mô hình: kỳ nào DropOut yesno? (1,2,3,4,5)

9/10/2011

Các kết quả mới: Thêm cột ”Bằng tốt nghiệp cấp 3”

