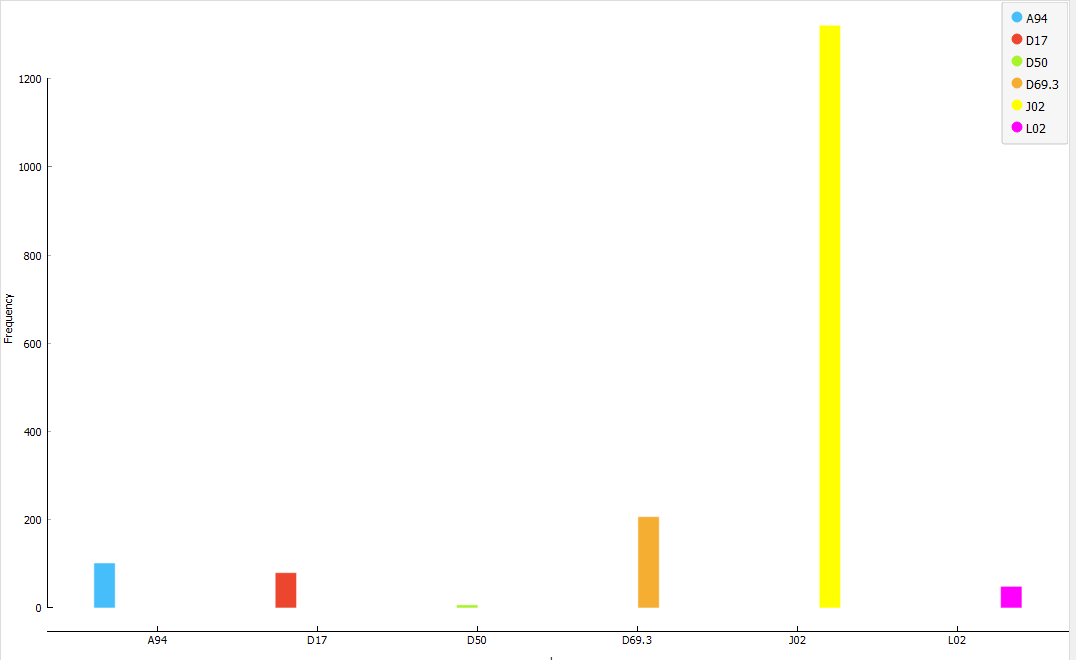
Kết quả thực nghiệm:

**Phân tích dữ liệu đầy đủ cả nam và nữ**

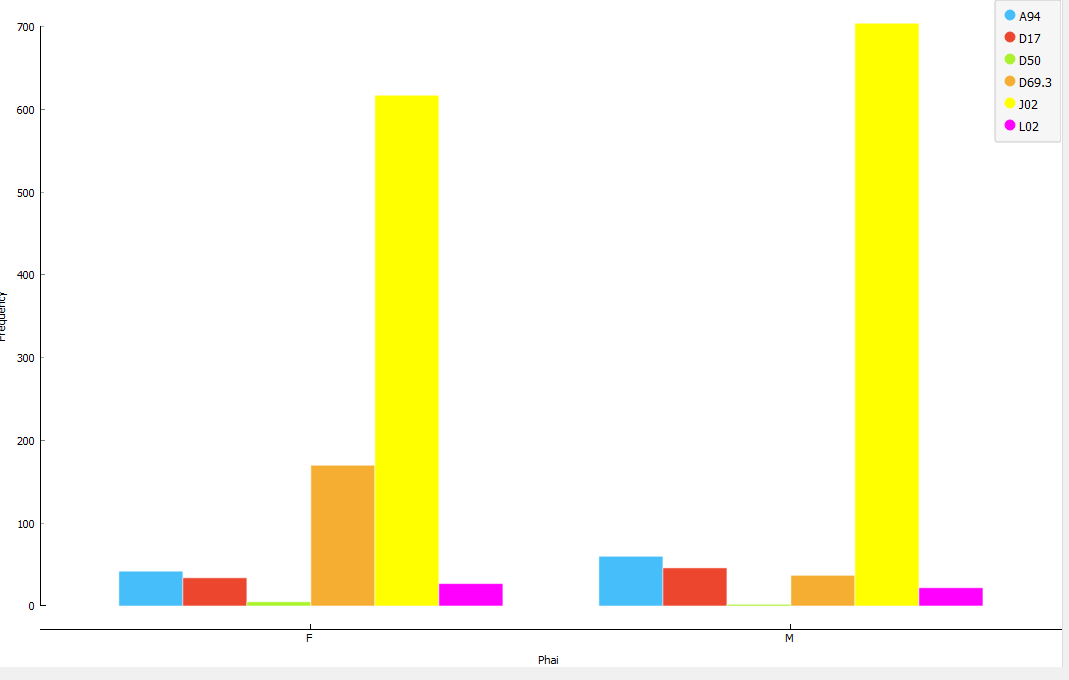
Thông tin dữ liệu đầy đủ

* 1766 mẫu
* 15 đặc trưng
* 6 lớp

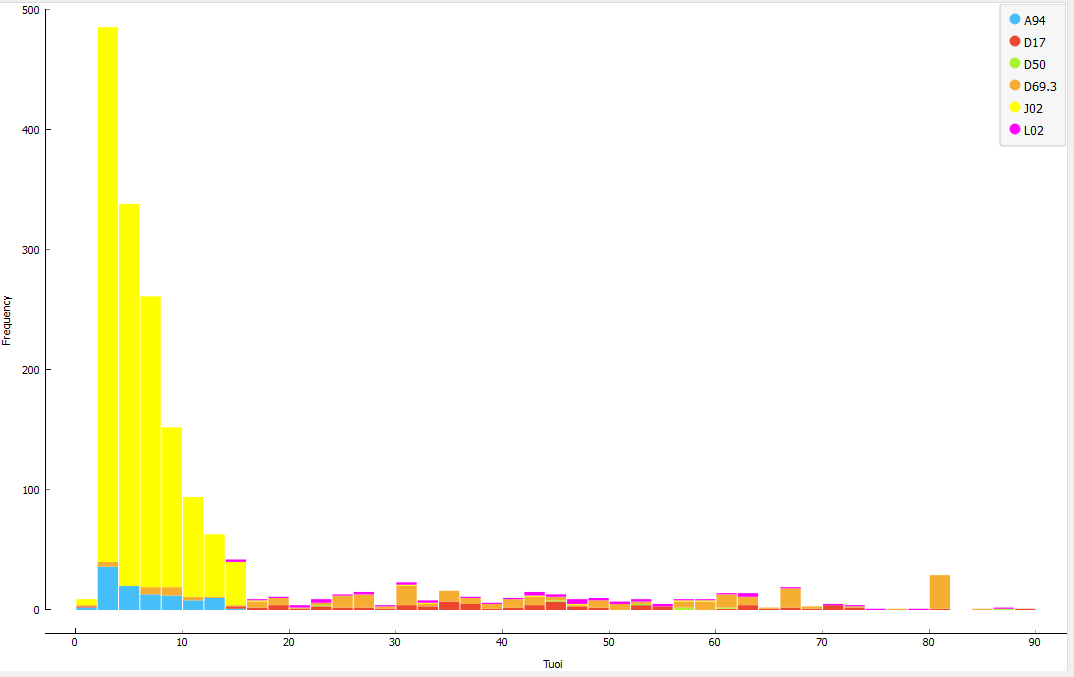
Phân bổ dữ liệu theo bệnh



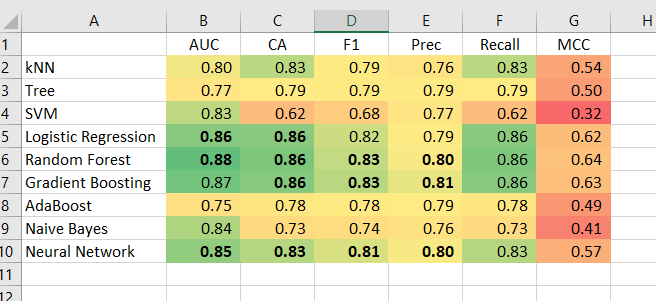
Phân bổ dữ liệu theo bệnh và giới tính



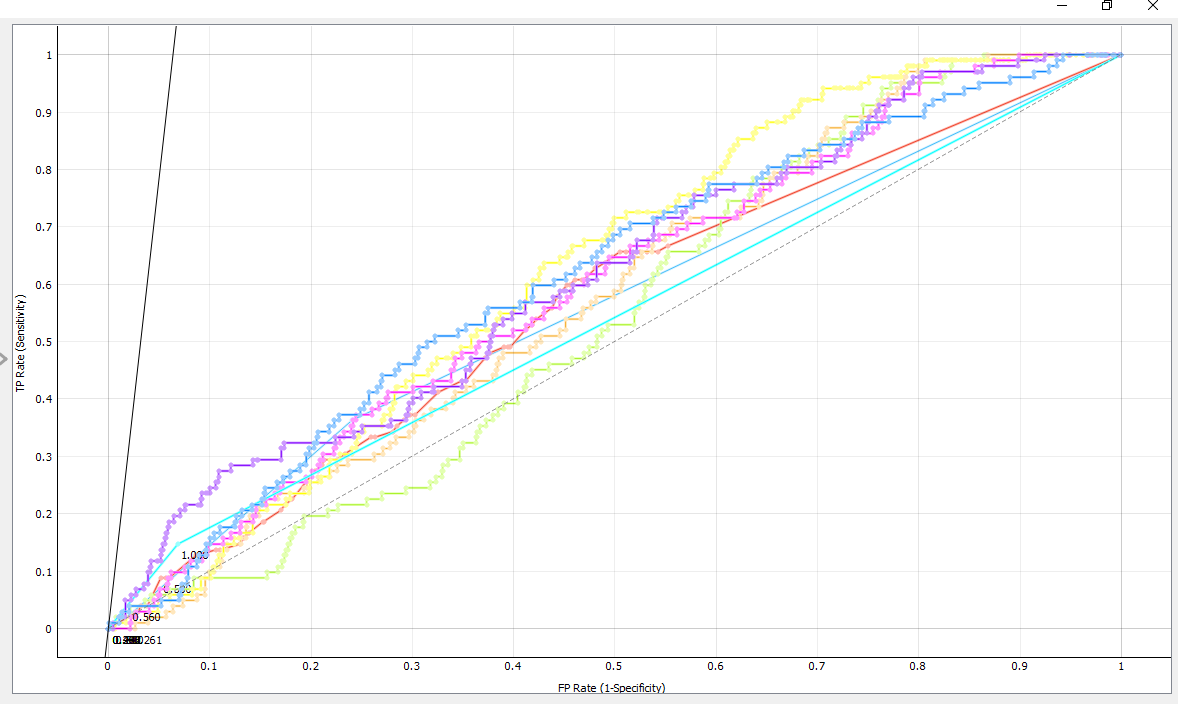
Phân bổ dữ liệu theo tuổi



Kết quả phân lớp



Kết quả phân tích ROC



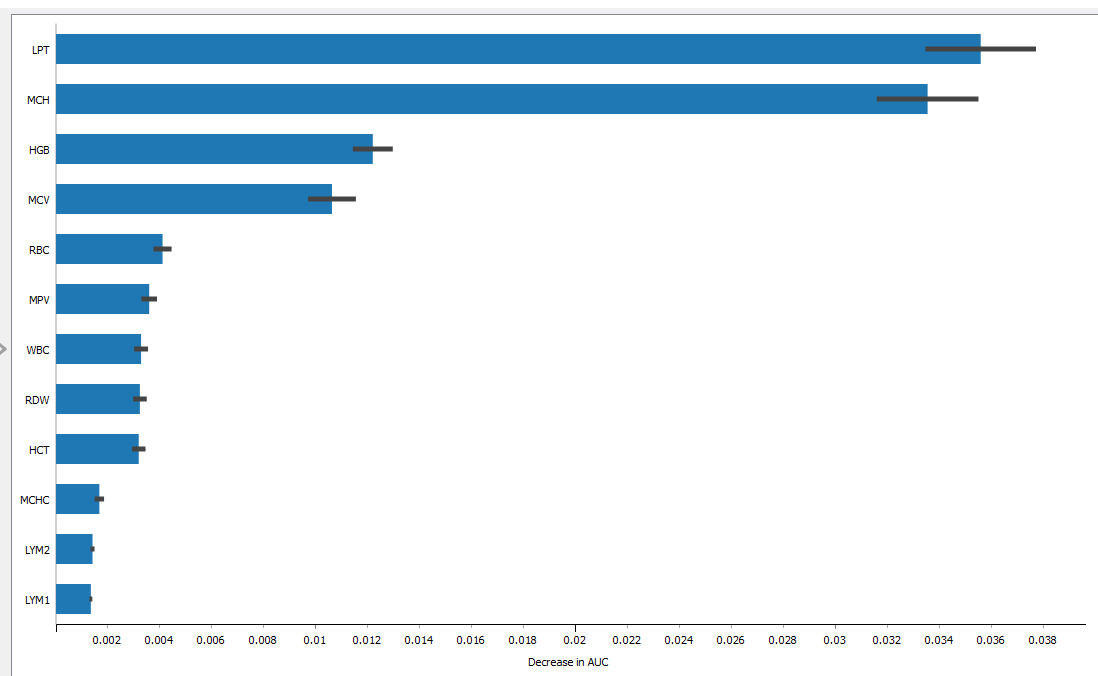
**So sánh với chuẩn hóa (Chuẩn hóa có kết quả tốt hơn 1 chút)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | AUC | CA | F1 | Prec | Recall | MCC |
| kNN | 0.81 | 0.84 | 0.80 | 0.78 | 0.84 | 0.56 |
| Tree | 0.77 | 0.79 | 0.79 | 0.79 | 0.79 | 0.50 |
| SVM | 0.83 | 0.62 | 0.68 | 0.77 | 0.62 | 0.32 |
| Logistic Regression | **0.86** | **0.86** | **0.82** | 0.79 | **0.86** | 0.62 |
| Random Forest | **0.88** | **0.86** | **0.83** | 0.80 | **0.86** | 0.64 |
| Gradient Boosting | **0.87** | **0.86** | **0.83** | 0.81 | **0.86** | 0.63 |
| AdaBoost | 0.75 | 0.78 | 0.78 | 0.79 | 0.78 | 0.48 |
| Neural Network | **0.85** | **0.83** | **0.81** | 0.80 | **0.83** | 0.57 |

**Thống kê đặc trưng**



**Sử dụng RF**

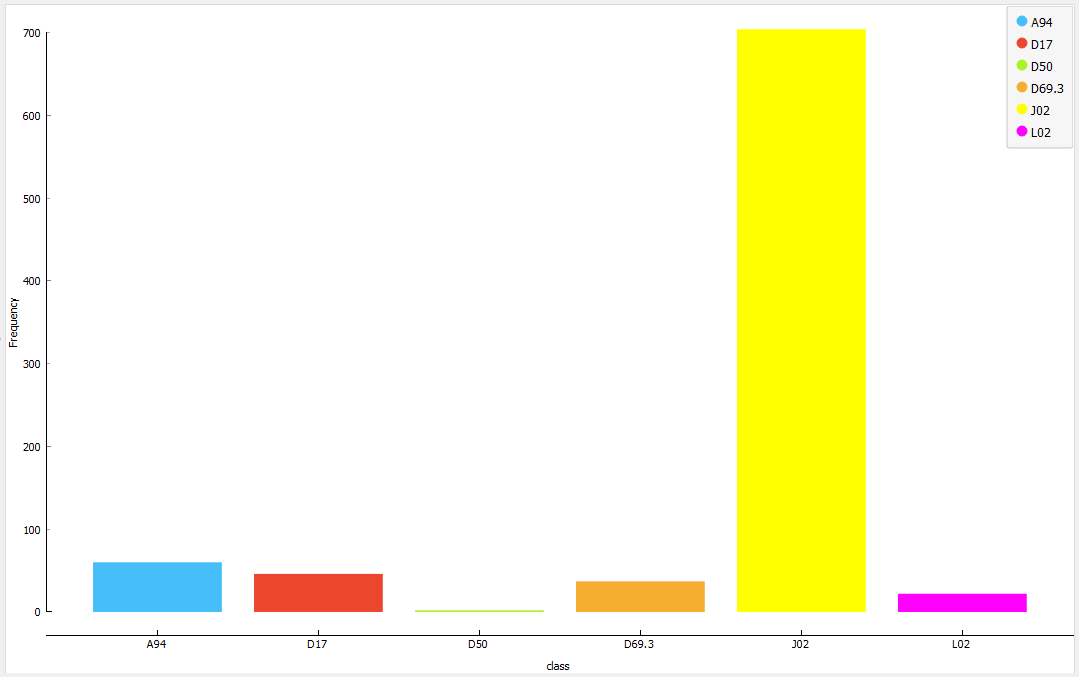


**Phân tích dữ liệu với bệnh nhân Nam**

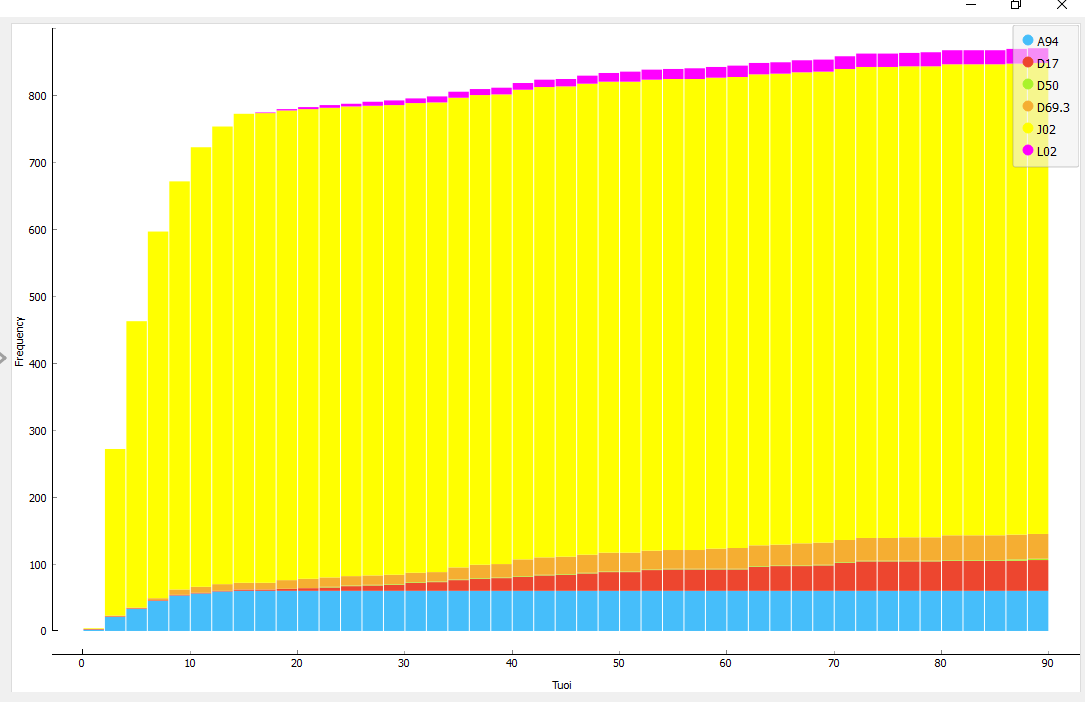
Tổng số mẫu: 871

Số đặc trưng: 15 đặc trưng

Phân bổ dữ liệu theo bệnh



Phân bổ dữ liệu theo tuổi

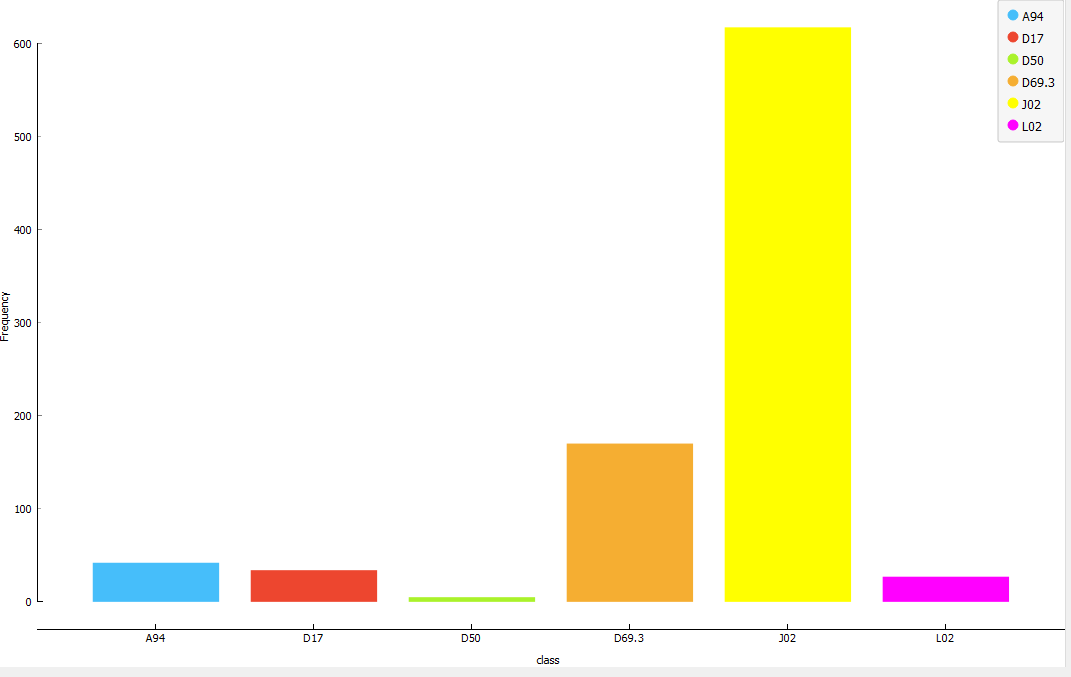


**Phân tích dữ liệu với bệnh nhân Nữ**

Thông tin dữ liệu đầy đủ

* 895 mẫu
* 15 đặc trưng
* 6 lớp

Phân bổ dữ liệu theo lớp



Phân bổ dữ liệu theo tuổi

