# Hybrid Approach to Identifying the Most Predictive and Discriminant Features in Supervised Classification Problems

Alexandre Bazin, Miguel Couceiro,
Marie-Dominique Devignes, Amedeo Napoli

06/07/2021

# The Problem

| | $gene\ 1$ | ... | $gene\ n$ | $Ill$ |
|---|---|---|---|---|
| $Patient\ 1$ | | | | Yes |
| ... | | Numbers ! | | ... |
| $Patient\ m$ | | | | No |

**Question**: Which features can best be used to predict (the development of) the disease ?

The answer constitutes a form of explanation of the supervised classification problem / dataset

# The Problem

"Predictive" "Discriminant"

| | $gene\ 1$ | ... | $gene\ n$ | $Ill$ |
|---|---|---|---|---|
| $Patient\ 1$ | | | | Yes |
| ... | | Numbers ! | | ... |
| $Patient\ m$ | | | | No |

# Prediction & Discrimination

### To predict

To assert that something will happen, is true.

⇒ A feature is said to be predictive when its value can be used to assert that an individual belongs to a particular class.

### To discriminate

To be able to perceive the differences between two things.

⇒ A feature is said to be discriminant when its value can be used to differentiate between the classes.

# Performance Measures

| | | | |
|---|---|---|---|
| True Positive (TP) | False Positive (FP) | Precision = $\frac{TP}{TP+FP}$ | $FDR = \frac{FP}{TP+FP}$ |
| False Negative (FN) | True Negative (TN) | FOR = $\frac{FN}{FN+TN}$ | $NPV = \frac{TN}{FN+TN}$ |
| Sensitivity = $\frac{TP}{TP+FN}$ | FPR = $\frac{FP}{FP+TN}$ | FScore = $2\frac{Precision \times Sensitivity}{Precision + Sensitivity}$ | |
| FNR = $\frac{FN}{TP+FN}$ | Specificity = $\frac{TN}{FP+TN}$ | Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$ | |
| Positive Likelihood Ratio = $\frac{Sensitivity}{FPR}$ | | MCC = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ | |
| Negative Likelihood Ratio = $\frac{FNR}{Specificity}$ | | | |

# Performance Measures

| | | | |
|---|---|---|---|
| True Positive (TP) | False Positive (FP) | Precision = $\frac{TP}{TP+FP}$ | $FDR = \frac{FP}{TP+FP}$ |
| False Negative (FN) | True Negative (TN) | FOR = $\frac{FN}{FN+TN}$ | $NPV = \frac{TN}{FN+TN}$ |
| Sensitivity = $\frac{TP}{TP+FN}$ | FPR = $\frac{FP}{FP+TN}$ | FScore = $2\frac{Precision \times Sensitivity}{Precision + Sensitivity}$ | |
| FNR = $\frac{FN}{TP+FN}$ | Specificity = $\frac{TN}{FP+TN}$ | Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$ | |
| Positive Likelihood Ratio = $\frac{Sensitivity}{FPR}$ | | MCC = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ | |
| Negative Likelihood Ratio = $\frac{FNR}{Specificity}$ | | | |

Sensitivity (Recall) = $\frac{TP}{TP+FN}$

$\Rightarrow$ Prediction (of class 1)

# Performance Measures

| | | | |
|---|---|---|---|
| True Positive (TP) | False Positive (FP) | Precision = $\frac{TP}{TP+FP}$ | $FDR = \frac{FP}{TP+FP}$ |
| False Negative (FN) | True Negative (TN) | FOR = $\frac{FN}{FN+TN}$ | $NPV = \frac{TN}{FN+TN}$ |
| Sensitivity = $\frac{TP}{TP+FN}$ | FPR = $\frac{FP}{FP+TN}$ | FScore = $2\frac{Precision \times Sensitivity}{Precision + Sensitivity}$ | |
| FNR = $\frac{FN}{TP+FN}$ | Specificity = $\frac{TN}{FP+TN}$ | Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$ | |
| Positive Likelihood Ratio = $\frac{Sensitivity}{FPR}$ | | MCC = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ | |
| Negative Likelihood Ratio = $\frac{FNR}{Specificity}$ | | | |

Specificity = $\frac{TN}{FP+TN}$

$\Rightarrow$ Prediction (of class 0)

# Performance Measures

| | | | |
|---|---|---|---|
| True Positive (TP) | False Positive (FP) | Precision = $\frac{TP}{TP+FP}$ | $FDR = \frac{FP}{TP+FP}$ |
| False Negative (FN) | True Negative (TN) | FOR = $\frac{FN}{FN+TN}$ | $NPV = \frac{TN}{FN+TN}$ |
| Sensitivity = $\frac{TP}{TP+FN}$ | FPR = $\frac{FP}{FP+TN}$ | FScore = $2\frac{Precision \times Sensitivity}{Precision + Sensitivity}$ | |
| FNR = $\frac{FN}{TP+FN}$ | Specificity = $\frac{TN}{FP+TN}$ | Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$ | |
| Positive Likelihood Ratio = $\frac{Sensitivity}{FPR}$ | | MCC = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ | |
| Negative Likelihood Ratio = $\frac{FNR}{Specificity}$ | | | |

Precision = $\frac{TP}{TP+FP}$

$\Rightarrow$ Correctness

# Performance Measures

| | | | |
|---|---|---|---|
| True Positive (TP) | False Positive (FP) | Precision = $\frac{TP}{TP+FP}$ | $FDR = \frac{FP}{TP+FP}$ |
| False Negative (FN) | True Negative (TN) | FOR = $\frac{FN}{FN+TN}$ | $NPV = \frac{TN}{FN+TN}$ |
| Sensitivity = $\frac{TP}{TP+FN}$ | FPR = $\frac{FP}{FP+TN}$ | FScore = $2\frac{Precision \times Sensitivity}{Precision+Sensitivity}$ | |
| FNR = $\frac{FN}{TP+FN}$ | Specificity = $\frac{TN}{FP+TN}$ | Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$ | |
| Positive Likelihood Ratio = $\frac{Sensitivity}{FPR}$ | | MCC = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ | |
| Negative Likelihood Ratio = $\frac{FNR}{Specificity}$ | | | |

Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$

$\Rightarrow$ Discrimination

# Performance Measures

| | | | |
|---|---|---|---|
| True Positive (TP) | False Positive (FP) | Precision= $\frac{TP}{TP+FP}$ | $FDR = \frac{FP}{TP+FP}$ |
| False Negative (FN) | True Negative (TN) | FOR= $\frac{FN}{FN+TN}$ | $NPV = \frac{TN}{FN+TN}$ |
| Sensitivity= $\frac{TP}{TP+FN}$ | FPR = $\frac{FP}{FP+TN}$ | FScore= $2\frac{Precision \times Sensitivity}{Precision + Sensitivity}$ | |
| FNR = $\frac{FN}{TP+FN}$ | Specificity = $\frac{TN}{FP+TN}$ | Accuracy= $\frac{TP+TN}{TP+TN+FP+FN}$ | |
| Positive Likelihood Ratio= $\frac{Sensitivity}{FPR}$ | | MCC = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ | |
| Negative Likelihood Ratio= $\frac{FNR}{Specificity}$ | | | |

Prediction Discrimination Correctness

# Permutation Importance of Features

Let
$T$ be a test set,
$f$ be a feature,
$M$ be a model,
$m$ be a measure.

$$impact(f, M, m) \approx \sum_{i=1}^{k} \frac{m(M, T_i^f) - m(M, T)}{k}$$

The feature $f$ is $\begin{cases} \text{predictive} \\ \text{discriminant} \end{cases}$, according to the model $M$,

if it has a negative impact on measures of $\begin{cases} \text{prediction} \\ \text{discrimination} \end{cases}$

# The Approach

**Two steps**:

- Identify the most predictive and/or discriminant features (machine learning + multicriteria decision making)
- Interpret and present their role in the problem according to the background knowledge on measures (multicriteria decision making + pattern mining)

# Identifying Important Features

Let $M$ be a model and $f_1, \ldots, f_5$ be five features

$$\begin{array}{ll} Accuracy: & f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\ Sensitivity: & f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\ Specificity: & f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \end{array}$$

# Identifying Important Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$
\begin{aligned}
c_1 = (RF, Accuracy) : &\quad f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 = (RF, Sensitivity) : &\quad f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 = (RF, Specificity) : &\quad f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 = (NN, Accuracy) : &\quad f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 = (NN, Sensitivity) : &\quad f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 = (NN, Specificity) : &\quad f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}
$$

# Identifying Important Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$\begin{aligned}
c_1 &= (RF, Accuracy): & f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 &= (RF, Sensitivity): & f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 &= (RF, Specificity): & f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 &= (NN, Accuracy): & f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 &= (NN, Sensitivity): & f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 &= (NN, Specificity): & f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}$$

**Important features** = Pareto front of this multicriteria decision problem

# Identifying Important Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$
\begin{aligned}
c_1 = (RF, Accuracy): &\quad f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 = (RF, Sensitivity): &\quad f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 = (RF, Specificity): &\quad f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 = (NN, Accuracy): &\quad f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 = (NN, Sensitivity): &\quad f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 = (NN, Specificity): &\quad f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}
$$

**Important features** = $\{f_4\}$

# Identifying Important Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$\begin{aligned}
c_1 &= (RF, Accuracy): & f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 &= (RF, Sensitivity): & {\color{red}f_2} \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 &= (RF, Specificity): & {\color{red}f_2} \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 &= (NN, Accuracy): & f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 &= (NN, Sensitivity): & f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 &= (NN, Specificity): & f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}$$

**Important features** = $\{{\color{red}f_2}, f_4\}$

# Identifying Important Features

Let $\{RF, NN\}$ be two models and $f_1, \dots, f_5$ be five features

$$
\begin{aligned}
c_1 &= (RF, Accuracy): & f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 &= (RF, Sensitivity): & f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 &= (RF, Specificity): & f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 &= (NN, Accuracy): & f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 &= (NN, Sensitivity): & f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 &= (NN, Specificity): & f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}
$$

**Important features** = $\{f_2, f_4, f_5\}$

# Identifying Important Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$\begin{aligned}
c_1 &= (RF, Accuracy): & f_4 \succ {\color{red}f_1} \succ f_2 \succ f_5 \succ f_3 \\
c_2 &= (RF, Sensitivity): & f_2 \succ {\color{red}f_1} \succ f_5 \succ f_3 \succ f_4 \\
c_3 &= (RF, Specificity): & f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 &= (NN, Accuracy): & f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 &= (NN, Sensitivity): & f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 &= (NN, Specificity): & f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}$$

**Important features** = $\{{\color{red}f_1}, f_2, f_4, f_5\}$

# Interpreting the Importance of Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$
\begin{aligned}
c_1 = (RF, Accuracy): &\quad f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 = (RF, Sensitivity): &\quad f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 = (RF, Specificity): &\quad f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 = (NN, Accuracy): &\quad f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 = (NN, Sensitivity): &\quad f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 = (NN, Specificity): &\quad f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}
$$

$f_2$ is important **because of** $c_2$ and $c_3$

# Interpreting the Importance of Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$
\begin{array}{ll}
c_1 = (RF, Accuracy): & f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 = (RF, Sensitivity): & f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 = (RF, Specificity): & f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 = (NN, Accuracy): & f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 = (NN, Sensitivity): & f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 = (NN, Specificity): & f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{array}
$$

$f_1$ is important **because of** $\{c_1, c_2\}$ (and others)

# Interpreting the Importance of Features

- $f_i$ is important because of Accuracy $\rightarrow$ $f_i$ is important for Discrimination
- $f_i$ is important because of Sensitivity $\rightarrow$ $f_i$ is important for Prediction
- $f_i$ is important because of $\{$Sensitivity, Specificity$\}$ $\rightarrow$ $f_i$ is important for Prediction
- $f_i$ is important because of $\{$Accuracy, Specificity$\}$ $\rightarrow$ $f_i$ is important for nothing in particular (?)

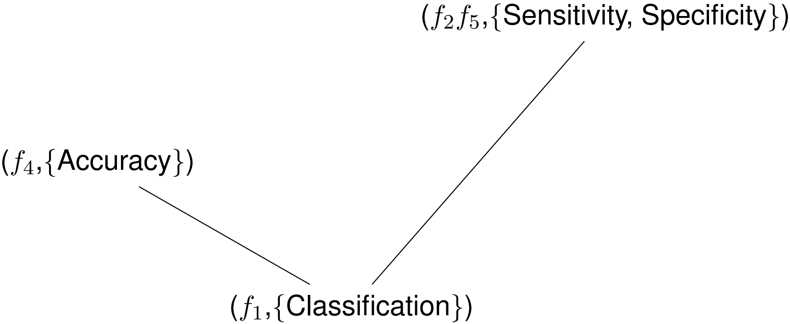
# Interpreting the Importance of Features

Let $\{RF, NN\}$ be two models and $f_1, \ldots, f_5$ be five features

$$\begin{aligned}
c_1 &= (RF, Accuracy): & f_4 \succ f_1 \succ f_2 \succ f_5 \succ f_3 \\
c_2 &= (RF, Sensitivity): & f_2 \succ f_1 \succ f_5 \succ f_3 \succ f_4 \\
c_3 &= (RF, Specificity): & f_2 \succ f_1 \succ f_3 \succ f_4 \succ f_5 \\
c_4 &= (NN, Accuracy): & f_4 \succ f_2 \succ f_1 \succ f_5 \succ f_3 \\
c_5 &= (NN, Sensitivity): & f_5 \succ f_1 \succ f_2 \succ f_3 \succ f_4 \\
c_6 &= (NN, Specificity): & f_5 \succ f_1 \succ f_2 \succ f_4 \succ f_3
\end{aligned}$$

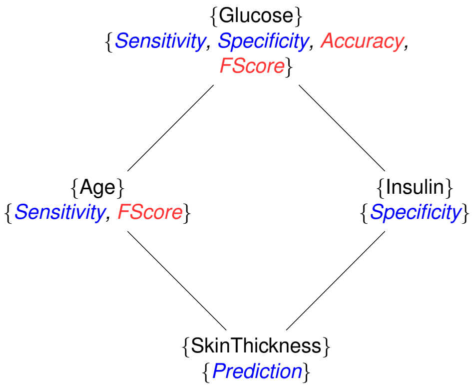
| | Accuracy | Sensitivity | Specificity | Prediction | Discrimination | Classification |
|---|---|---|---|---|---|---|
| $f_1$ | | | | | | $\times$ |
| $f_2$ | | $\times$ | $\times$ | $\times$ | | $\times$ |
| $f_4$ | $\times$ | | | | $\times$ | $\times$ |
| $f_5$ | | $\times$ | $\times$ | $\times$ | | $\times$ |

# Interpreting the Importance of Features



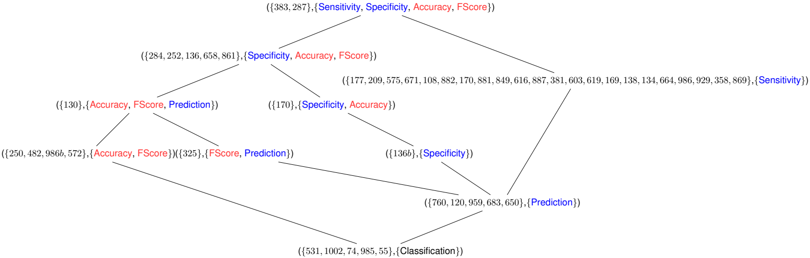
# Example

768 instances, 8 features
4 important ones



# Example

111 instances, 1195 features
47 important ones



**Merci !**