Introduction to Machine Learning
NPFL 054

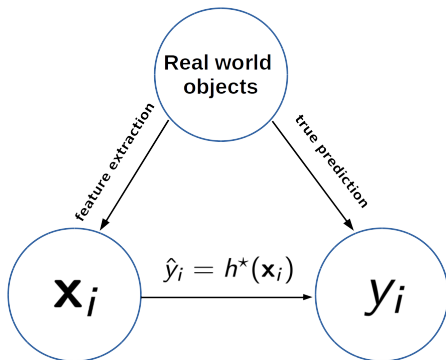`http://ufal.mff.cuni.cz/course/npfl054`

Barbora Hladká
hladka@ufal.mff.cuni.cz

Martin Holub
holub@ufal.mff.cuni.cz

Charles University,
Faculty of Mathematics and Physics,
Institute of Formal and Applied Linguistics

Imagine that you are able to develop a really optimal classifer. Is the zero test error always feasible?

Bayes classifier and Bayes error

**Imagine that you are able to develop a really optimal classifer.
Is the zero test error always feasible?**

> The **Bayes classifier** minimises the probability of misclassification

Thus, by definition, error produced by the Bayes classifier is irreducible and is called *Bayes error*.

What is the lowest possible error rate

Bayes classifier assigns each example to the most likely class, given its feature values

$$\hat{y} = max_y \Pr(y \mid \mathbf{x})$$

The Bayes classifier produces the lowest possible test error rate,
so called **Bayes error rate**

$$1 - \mathsf{E}\left(max_y \Pr(y \mid \mathbf{x})\right)$$

What is the lowest possible error rate

Practical view on your development data

Are there identical feature vectors in your data set?

- Get the same feature vectors
- How many of them have the same target value?

Examination requirements

- Bayes classifier and Bayes error – definition and meaning