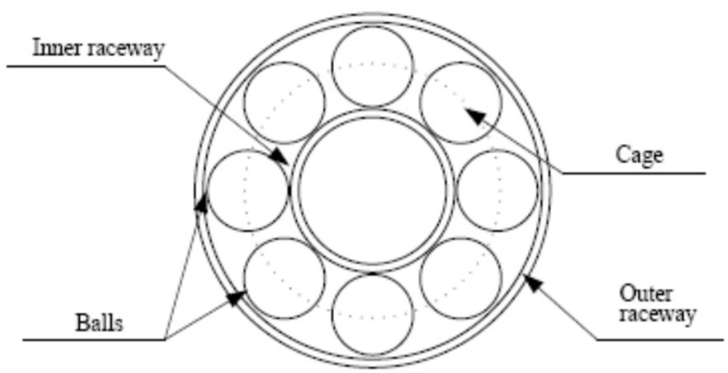
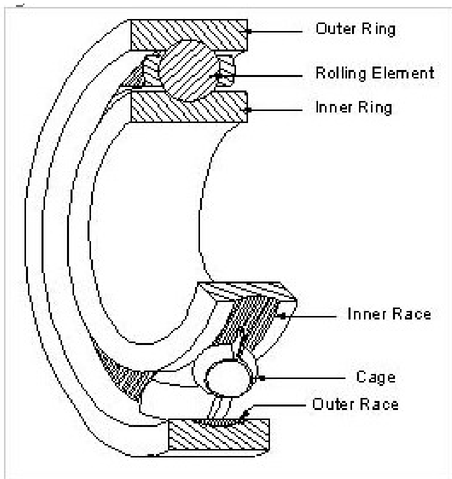
# Application Example

***Beatrice Lazzerini***
Department of Information Engineering
University of Pisa
ITALY

# Condition-based maintenance

- Condition-based maintenance (CBM) aims to detect the early occurrence and severity of a fault and to identify faulty components.
- We use multiple sensors for CBM of rotating machines, the most critical components of which are bearings.



- To perform early detection and diagnosis of bearings, we need to consider three main issues:

VANNO SCELTI SEMPRE

  - the choice of the representation space – ESEMPIO: COME RAPPRESENTARE IL SEGNALE FISICO.
  - the feature selection algorithm
  - the classifier

# Representation space

- **time-domain analysis**
  - performance indexes such as Crest Factor, Root Mean Square, Kurtosis, etc.
- **frequency-domain analysis**
  - spectrum analysis by means of the Fast Fourier Transform (FFT)
- **time-frequency domain analysis**

# Feature selection

- Several different feature selection algorithms:
  - **individual feature selection**
  - **forward feature selection**
  - **backward feature selection**
  - ….

# Classifier

The choice of the classifier depends on several factors, such as:

- the application domain
- the number and type of features
- user preferences

# Experiments

- We used vibration signals coming from a mechanical device that includes more than ten rolling element bearings monitored, in the time domain, by four accelerometers.
- The signals were collected both with all faultless bearings and after the replacement of a faultless bearing with a damaged one.
- The bearings were artificially damaged.
- Four types of damages were considered so that the data can be classified into five classes:
  - C1: faultless bearing
  - C2: indentation on the inner raceway (450 μm indentation)
  - C3: indentation on the roll
  - C4: sandblasting of the inner raceway
  - C5: unbalanced cage
- The fault of class C3 can be divided into three subclasses depending on the severity of the damage:
  - C3.1: 450 μm indentation on the roll (light)
  - C3.2: 1.1 mm indentation on the roll (medium)
  - C3.3: 1.29 mm indentation on the roll (high)

| Class | Number of signals (sec) |
|---|---|
| C1 | 2890 |
| C2 | 1770 |
| C3 | 4790 |
| C4 | 1520 |
| C5 | 1770 |

| Class | Number of signals (sec) |
|---|---|
| C3.1 | 1770 |
| C3.2 | 1250 |
| C3.3 | 1770 |

- We used frequency-domain analysis by transforming the time vibration signals by means of FFT.
- Based on heuristic considerations, for each accelerometer, we considered the frequency interval [1, 300] Hz, sampled every 1 Hz.
- Each frequency sample represents a *feature*.
- Therefore, if we consider a single accelerometer, each signal can be represented in $\Re^n$ with $n \leq 300$. If we consider the four accelerometers, each signal can be represented with a maximum of $n=300\times4=1200$ features.
- In each experiment we adopt the forward feature selection.
- The training is performed using **balanced** data: we use the **random undersampling** technique so that all classes contain the same number of samples as the least numerous one.
- The training set is built using the *hold-out* method, randomly choosing 70% of the data (the remaining 30% are used as test set).

- Classifiers used:
  - LDC (Linear discriminant classifier)
  - QDC (Quadratic discriminant classifier)
  - MLP (Multi-Layer Perceptron)
- LDC and QDC are used to perform both feature selection and signal classification, while MLPs perform classification of the signals represented by the features selected by LDC and QDC.

# First experiment

- We aim to classify the signals into **two classes**: faultless bearings (C1) and damaged bearings (C2, C3, C4, C5).
- For each accelerometer, we perform forward feature selection to find the best *discriminating frequencies* (DFs).
- In this way, in addition to decreasing the space dimension for signal representation, we also identify the most significant accelerometers.
- We use one LDC and one QDC for each accelerometer.
- We repeat the experiment 10 times for each accelerometer (limiting the maximum number of selected features to 10).
- Then, for each accelerometer and each classifier, we select the best *m* ($m \leq 10$) features which are exactly the same (and in the same order) for all the trials. We call these features *stable features* (SFs).
- Using the stable features, we calculate the accuracy on 10 more trials, then we compare the four accelerometers based on the classification accuracy, expressed in the form (mean ± standard deviation).

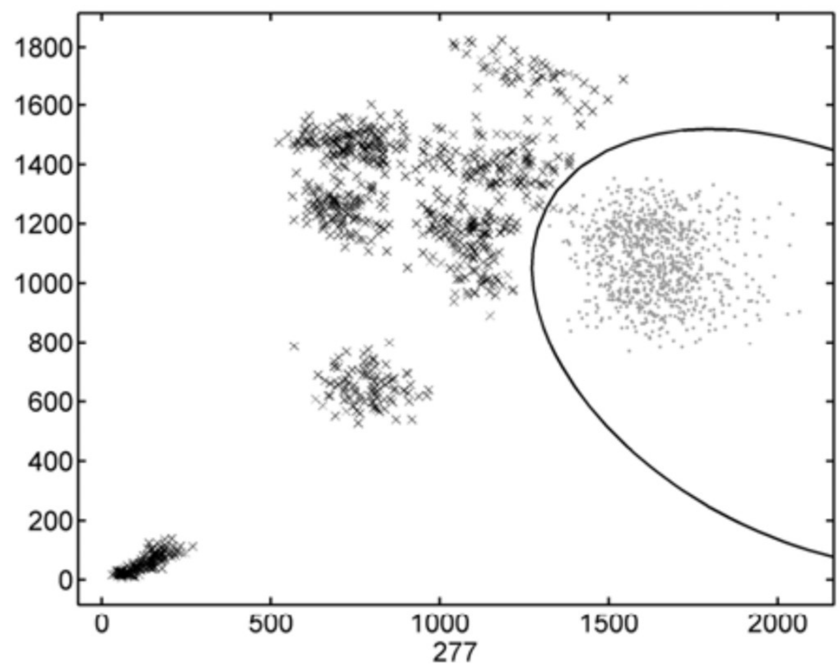
Classification of C1, C6

| Accelerometer | LDC | | QDC | |
|---|---|---|---|---|
| | Number of SFs | Accuracy Mean±Std.Dev | Number of SFs | Accuracy Mean±Std.Dev |
| 1 | 1 | 96.04±0.97 % | 1 | 98.16±0.42 % |
| 2 | 1 | 85.43±0.68 % | 0 | N.A. |
| 3 | 1 | 94.77±0.43 % | **2** | **99.69±0.14 %** |
| 4 | 1 | 92.21±0.68 % | 2 | 96.63±0.25 % |

C6 = {C2,C3,C4,C5}, i.e., set of all damaged bearings

NA = no SFs have been selected

- QDC applied to the features of the third accelerometer achieves the best accuracy with a very small standard deviation.
- We have reduced the feature space from $\Re^{300}$ to $\Re^{2}$.




Separation of C1 (gray dots) and C6 (black crosses) for a test set produced by QDC using the SFs of the third accelerometer.

# Second experiment

- The goal is to classify the signals into **five classes** C1, C2, C3, C4, C5 in order to recognize the different types of faults regardless of the severity of the fault.
- Again, we repeat the experiment 10 times for each accelerometer, using LDC and QDC.

Classification of C1, C2, C3, C4, C5

| Accelerometer | LDC | | QDC | |
|---|---|---|---|---|
| | Number of SFs | Accuracy Mean±Std.Dev | Number of SFs | Accuracy Mean±Std.Dev |
| 1 | 0 | N.A. | 0 | N.A. |
| 2 | **3** | **98.26±0.27 %** | 2 | 96.66±0.28 % |
| 3 | **3** | **96.87±0.36 %** | 2 | 94.92±0.38 % |
| 4 | 3 | 95.85±0.58 % | 3 | 96.69±0.37 % |

- The second and third accelerometers are the best, but *we would like to achieve higher classification accuracy*. We decide, on the basis of a heuristic reasoning, to consider "good" an accuracy of at least 99.00 % and a standard deviation close to 0.10 %. This means we want a *good stable* classifier.
- ➔ *let us use classifier fusion*

- We fuse
  - LDC, Accelerometer 2
  - LDC, Accelerometer 3

  using *max*, *min* and *mean*

| | Accuracy Mean±Std.Dev |
|---|---|
| Combiner (MAX) | 98.88±0.16 % |
| Combiner (MIN) | 99.47±0.21 % |
| Combiner (MEAN) | 99.10±0.16 % |

- As shown in the table, no combiner respects the fixed accuracy thresholds.

  ➔ we combine <u>four</u> classifiers, namely the previous <u>LDC</u>s, and <u>two</u> MLPs applied, respectively, to the features selected by the two LDCs.
- The two MLPs are the most accurate of several different architectures (different number of hidden neurons, neuron activation function, number of epochs, etc.).

- We obtain

| | Accuracy Mean±Std.Dev |
|---|---|
| **Combiner (MAX)** | 98.93±0.15 % |
| **Combiner (MIN)** | **99.42±0.08 %** |
| **Combiner (MEAN)** | **99.32±0.09 %** |

- In practice, we behave as follows
  - we start using statistical classifiers and, if the accuracy thresholds are not met, we first resort to the combination of these statistical classifiers, then we combine the statistical classifiers with neural networks.

# Third experiment

- This experiment aims to classify the signals into seven classes C1, C2, C3.1, C3.2, C3.3, C4, C5 in order to recognize both the different types of faults and the different levels of severity.
- By combining six classifiers
  - one QDC (which selects 2 stable features of accelerometer 3)
  - one LDC (4 SFs of accelerometer 2)
  - one LDC (3 SFs of accelerometer 3)
  - three MLPs applied, respectively, to the features selected by the previous classifiers

  we obtain that *max* and *mean* respect the chosen accuracy thresholds.

| | |
|---|---|
| Combiner (MAX) | 99.27±0.13% |
| Combiner (MEAN) | 99.06±0.12% |

# Noise analysis

- We trained neural classifiers (MLPs and RBF networks) with signals that have not undergone any alteration.
- We added a white Gaussian noise signal to all vibration signals of the test set in the time domain.
- We performed 100 trials. The noise generation was repeated for each trial and for each signal sample.
- For each trial we used an increasing noise level.
- We tried several different configurations for MLPs and RBF networks. RBF networks showed the best performance for lower noise levels, while MLPs proved to be the best classifiers for higher noise levels.