**Today's Schedule**

- **Topic revision**
- **Maths formulae to remember**
- **Methods/derivations to understand**
- **Exam technique**

**Topics dealt within the course**

- Distance and similarity measures (Pearson correlation coeff.)
- Clustering (K-means clustering)
- Dimensionality reduction (covariance matrix, PCA)
- Classification
  - K-NN classification
  - Naive Bayes
  - Gaussian classifiers (MLE, discriminant functions)
  - Neural networks (Perceptron error correction algorithm, sum-of-squares error cost function, gradient descent, EBP)
- Statistical pattern recognition theories
  - Bayes theorem, and Bayes decision rule
  - Probability distributions and parameter estimation
    - Bernoulli distribution / Multinomial distribution
    - Gaussian distribution
    - Discriminant functions
    - Decision boundaries/regions (minimum error rate classification)
  - Evaluation measures and methods
- Optimisation problems

---

**Maths formulae to remember**

- Univariate Gaussian pdf:
  \[ p(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right) \]

- Multivariate Gaussian pdf:
  \[ p(x | \mu, \Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp \left( -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right) \]

- Correlation coefficient:
  \[ \rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \sigma_{jj}}} \]

- Logistic sigmoid function:
  \[ y = g(a) = \frac{1}{1 + \exp(-a)} \]
  \[ g'(a) = g(a)(1-g(a)) \]

- Softmax activation function (for multiple output nodes):
  \[ y_k = \frac{\exp(a_k)}{\sum_{l=1}^{K} \exp(a_l)} \]

- and basic maths rules (e.g. differentiation)

---

**Maths formulae to remember (cont.)**

- **Bayes decision rule** (cf. MAP decision rule)
  \[ k^* = \arg \max_k P(C_k | x) = \arg \max_k P(x | C_k) P(C_k) \]

- **Naive Bayes for document classification**
  - Likelihood by Bernoulli document model
    \[ P(b(C_k)) = 1 - \prod_{n=1}^{N} (1 - p_w(C_k | x)) \]
  - Likelihood by Multinomial document model
    \[ P(C_k) \propto \prod_{n=1}^{N} P(w_n | C_k)^{||x||_n} = \prod_{n=1}^{N} P(w_n | C_k) \]

- **Discriminant functions of Gaussian Bayes classifiers**
  - Class 
  \[ k_n \mid x \}
  - **Evaluation measures and methods**
    - Decision boundaries/regions (minimum error rate classification)
    - Likelihood by Multinomial document model
      \[ P(C_k) \propto \prod_{n=1}^{N} P(w_n | C_k)^{||x||_n} = \prod_{n=1}^{N} P(w_n | C_k) \]
  - Parameter estimation from samples:
    - **Logistic sigmoid function**
      \[ y = g(a) = \frac{1}{1 + \exp(-a)} \]
      \[ g'(a) = g(a)(1-g(a)) \]
    - **Softmax activation function** (for multiple output nodes):
      \[ y_k = \frac{\exp(a_k)}{\sum_{l=1}^{K} \exp(a_l)} \]

---

**Methods/derivations to understand**

- Clustering and classification
  - Maximum likelihood estimation
  - Gradient descent and back propagation algorithm (neural networks) for minimising the sum-of-squares error

- Dimensionality reduction and feature selection (NE)

NB: Learning is a difficult problem by nature — generalisation from a limited amount of training samples. Need to assume some structures (constraints).
### Exam revision

Look at lecture notes, slides, tutorials, coursework, and past papers.

**Early exam papers:** many (useful) multiple choice Qs
- No longer the exam format
- Syllabus has changed slightly
- Recent exam papers since 2008/09
  - Answer two questions from section A (ADS) and two questions from section B (Learning).
  - Closed-book exam.
  - Calculators may be used (approved ones only).
  - Solutions are available only for 2008/09, 2009/10, 2013/14 (no plans of releasing those of missing years)
  - NB: errors in some solutions, e.g. 5 (c) of 2008/09: square root is not taken in computing standard deviations.
- Well prepared for the exam of 120 minutes
  - 60 minutes/section, 30 minutes/question

### Exam revision (cont.)

- **Don’t overfit!**
  - Anything that appears in the notes, slides, tutorial sheets, or coursework is examinable, unless marked non-examinable, extra topics, or †
- **Don’t trust unofficial solutions**

### Time in the exam

- **Half an hour per question** (minus time to pick questions)
- **Don’t panic!**
- **Go for easy marks first**
- **Don’t spend a long time on any small part**
- **Don’t scrawl - you might lose marks if the marker cannot read/understand**
- **Know the standard stuff:**
  - there’s not time to work everything out from scratch

Calculators may be used in the examination. The School of Informatics does not provide calculators for use in exams. If the use of a calculator is permitted in an exam, it’s your responsibility to bring an approved calculator to the exam.

### Inf2b Revision Meeting

- **Date:** TBC (in late April)
- **Send me questions/requests that you want me to discuss at the meeting.**

---

**End-of-course feedback:**

---

**Thanks!**