

## 1 Sobre os microprojetos

### 1.1 Objetivo

O objetivo do microprojeto é de realizar um trabalho de casa sobre a análise e implementação de um tema abordado nas aulas com uma parte de programação e de experiência numérica.

- Serão criados grupos de 5-6 alunos;
- Devem produzir um relatório de 6 páginas e um conjunto de slides para apresentar publicamente o vosso trabalho.
- A classificação deste trabalho será: 50% do guião, 50% da apresentação oral (marcada para a última semana de Maio).
- Cada grupo deve escolher uma proposta na lista apresentada adiante. Um tema não pode ser escolhido mais do que duas vezes.

### 1.2 Procedimentos

1. Até dia 10/5, formação dos grupos, onde cada grupo deve designar um representante que me enviará um email com a lista dos membros do grupo.
2. Dia 10/05, a partir de 17h00, cada representante de grupo envia-me um email com o nome do projeto que gostaria realizar. Vou atribuir os projetos consoante a hora de chegada dos emails (excluindo os emails que chegarem mais cedo).
3. Se um tema já foi escolhido duas vezes, o grupo deve escolher uma outra proposta.

## 2 Temas

Os temas, detalhados no resto do documento, são os seguintes.

- *T1 Uma versão dual do logistic classifier com Kernel*
- *T2 Operador de suavização (smooth operator) para acelerar a aprendizagem*
- *T3 Multi-logistic classifiers versus softmax*
- *T4 Método de amostragem dinâmica*
- *T5 Método SAGA*

### 2.1 Uma versão dual do classificador logístico

**Descrição.** Propõe-se neste trabalho a introdução da versão dual do logistic classifier do mesmo modo que no caso do perceptron. Analisamos os ganhos em tempo computacional assim que a eficiência de aprendizagem. A seguir, consideramos a versão Kernel do logistic classifier para classificar dados que não são linearmente separáveis. Utilizamos kernel de tipo polinomial ou exponencial.

**Trabalhos a realizar.**

1. descrição do método dual;

2. implementação e validação;
3. introdução do Kernel;
4. implementação e benchmarks.

Os documentos de referências são `documento1-1.pdf` e `documento1-2.pdf`.

## 2.2 Data dependent learning rate for logistic classifier

**Descrição.** No processo de otimização estocástico, a escolha do passo de aprendizagem é determinado com a posição relativa da amostragem e o hiperplano de separação. O objetivo é desenvolver um procedimento matematicamente justificado para avaliar o passo de maneira a otimizar a aprendizagem.

### Trabalhos a realizar.

1. Descrição matemática dos parâmetros de aprendizagem;
2. Implementação de várias técnicas;
3. Avaliação da qualidade do resultado e rapidez de aprendizagem.

Os documentos de referências são `documento2-1.pdf` e `documento2-2.pdf`.

## 2.3 Multi-logistic classifiers versus softmax

**Descrição.** Quando a classe é multinomial (mais de dois valores para a classe), podemos usar uma combinação de *logistics classifier* em lugar do *softmax classifier*. Neste trabalho estudamos vários cenários para substituir o softmax (um contra todos, um contra um). Pretende-se avaliar a eficiência desta metodologia ambos em termo de precisão que de rapidez de execução.

### Trabalhos a realizar.

1. Implementar as versões um contra todos e dicotômica;
2. Avaliação e comparação com o softmax;
3. Procedimento para determinar a melhor escolha dos valores de classe para classificar.

Os documentos de referências são `documento3-1.pdf` e `documento3-2.pdf`.

## 2.4 Método de amostragem dinâmica

**Descrição.** Os chamados métodos de redução de ruído (do inglês *noise reduction methods*) conseguem redução de ruído, aumentando gradualmente o tamanho do mini-batch usado no cálculo do gradiente estocástico, usando assim, estimativas do gradiente cada vez mais precisas à medida que o processo de otimização prossegue. Uma classe destes métodos é a dos métodos de amostragem dinâmica (do inglês *dynamic sampling methods*). Neste trabalho pretende-se implementar um método de amostragem dinâmica, e analisar o seu desempenho na resolução de um dos problemas de machine learning estudados nesta UC.

### Trabalhos a realizar.

1. Implementação de um método amostragem dinâmica;
2. Análise do desempenho em termos de eficiência e qualidade das soluções produzidas.
3. Análise dos desempenhos do método de amostragem dinâmica *versus* método do gradiente estocástico mini-batch.

Os documentos de referências são: [1] `documento4.pdf`, [2] L. Bottou, F. E. Curtis, J. Nocedal. Optimization Methods for Large-Scale Machine Learning. Northwestern University, 2018.

## 2.5 Método SAGA

**Descrição.** Os chamados métodos de redução de ruído (do inglês *noise reduction methods*) conseguem redução de ruído, aumentando gradualmente o tamanho do mini-batch usado no cálculo do gradiente estocástico, usando assim, estimativas do gradiente cada vez mais precisas à medida que o processo de otimização prossegue. Uma classe destes métodos é a dos métodos de agregação do gradiente (do inglês *gradient aggregation methods*). Neste trabalho pretende-se implementar o método SAGA, e analisar o seu desempenho na resolução de um dos problemas de machine learning estudados nesta UC.

### Trabalhos a realizar.

1. Implementação do método SAGA;
2. Análise do desempenho em termos de eficiência e qualidade das soluções produzidas.
3. Análise dos desempenhos do método SAGA *versus* método do gradiente estocático mini-batch.

Os documentos de referências são: [1] `documento4.pdf`, [2] L. Bottou, F. E. Curtis, J. Nocedal. Optimization Methods for Large-Scale Machine Learning. Northwestern University, 2018.