

1. P1 - Métodos de Classificação

1.1. Features

A partir da análise do ficheiro de treino que foi disponibilizado, foram escolhidas as seguintes features que permitiram atingir um erro de 0.0:

- Tamanho da palavra;
- Número de acentos nas palavras;
- Se a palavra tem um 'a' ou não.

Foram testadas outras *features*, contudo, não mostraram contribuições significantes para as considerarmos.

1.2. Classificador

O método de aprendizagem que foi utilizado como classificador foi o *Decision Tree Learning (DTL)*. Outros algoritmos de aprendizagem também foram testados sem ajustar parâmetros (*K-Nearest Neighbors* e *Stochastic Gradient Descent*) porém obtiveram resultados inferiores ao *DTL*. Desta forma, foi decidido que a melhor escolha era de facto o *DTL*. Não foi necessário ajustar parâmetros para obter o erro desejado uma vez que este já era mínimo.

2. P2 - Métodos de Regressão

2.1. Métodos Seleccionados

O primeiro método de regressão seleccionado foi o *Kernel Ridge (KR)*. O motivo da escolha deve-se ao facto do mesmo ser capaz de se ajustar melhor aos dados de ambos os testes e apresentar um erro dentro dos limites de tolerância. Foram necessários alterar parâmetros do *KR*, nomeadamente, o α para 0.001 (que neste caso ajusta a curva de forma mais suave aos dados), o γ para 0.10 (que ajusta a distância a que os pontos são considerados mais semelhantes) e o tipo de função kernel para a *Radial Basis (RBF)*. Os ajustes foram realizados de forma a obter um menor erro através do método de validação cruzada.

O segundo método de regressão escolhido foi o *Linear Regression (LR)* que apresenta um erro dentro dos limites de tolerância apenas no segundo ajuste. Este método consegue ajustar-se a um dos conjuntos de dados pois, neste caso em particular a alteração do parâmetro *fit_intercept* para *False* admite que os dados já estão centrados (o que não se verifica) contudo as características dos dados permitem que a reta se ajuste à média dos dados, na direcção correta como se pode verificar na figura 2..

Figura 1: Dataset 1

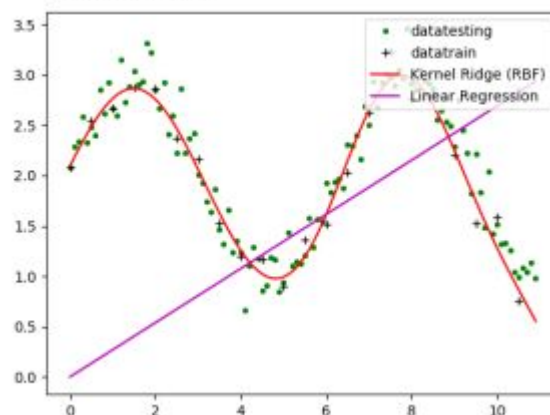
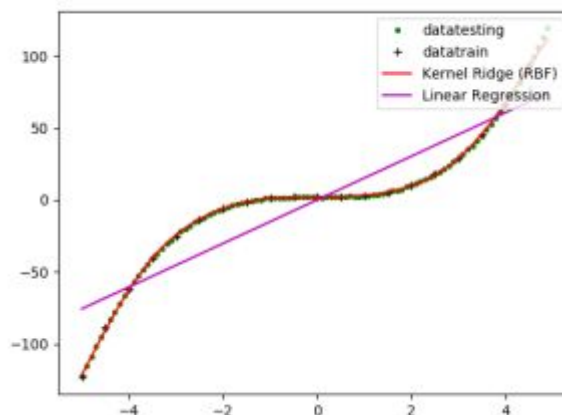


Figura 2: Dataset 2



3. P3 - Aprendizagem por Reforço

Foi observado que com uma learning rate mais lenta (α mais pequeno, ou seja, 0.01) o agente não conseguiu obter um erro nos valores Q dentro dos limites de tolerância do primeiro ambiente, contudo, com um learning rate mais rápido (α maior, ou seja, 0.1) ele obteve sucesso nos dois ambientes. Assim, conclui-se que o learning rate para o primeiro ambiente não pode ser tão baixo se quiser ter um Q dentro dos limites de tolerância

3.1. Representação Gráfica dos ambientes

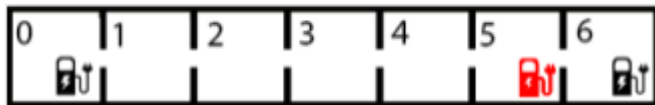
Através da análise das trajetórias foi possível perceber a forma dos dois ambientes nos quais o agente se move. Ambos os ambientes não são determinísticos pois no estado 5 existem duas opções para a mesma ação, o que justifica a escolha da política utilizada (política estocástica com a utilização da função softmax). O ambiente foi considerado como sendo uma série de salas ligadas por portas em que o agente se move. O estado objetivo foi representado

como uma sala na qual existe uma estação de carregamento para o *robot* específico à qual ele tem que se dirigir da forma mais rápida possível.

3.1.1. Primeiro Ambiente

O primeiro ambiente tem o formato de um corredor começando no estado 0 e acabando no estado 6, tal como se pode verificar na figura 1

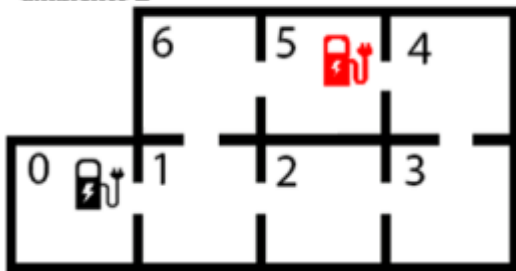
Figura 3: Representação gráfica do ambiente 1



3.1.2. Segundo Ambiente

O segundo ambiente tem o formato de um corredor “circular” entre o estado 1 e o estado 6, estando o estado objetivo (estado 0) ligado ao estado 1.

Figura 4: Representação gráfica do ambiente 2



3.2. Função de Recompensa

O valor das recompensas para cada par (estado, ação) pode ser obtido na respectiva entrada da tabela.

	Ação 0	Ação 1
Estado 0	1	1
Estado 1	0	0
Estado 2	0	0
Estado 3	0	0
Estado 4	0	0
Estado 5	0	0
Estado 6	1	1

3.3 Movimento do Agente

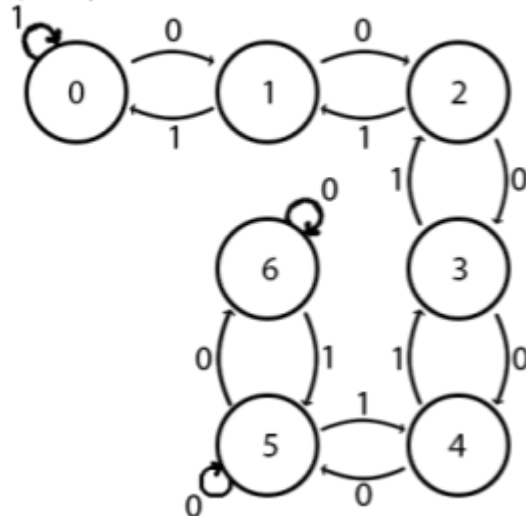
As ações que este agente suporta são:

- Ação 0: Ir para a próxima sala.

- Ação 1: Ir para a sala anterior.
- Ambos também podem representar “ficar na mesma sala”.

3.3.1. Primeiro Ambiente

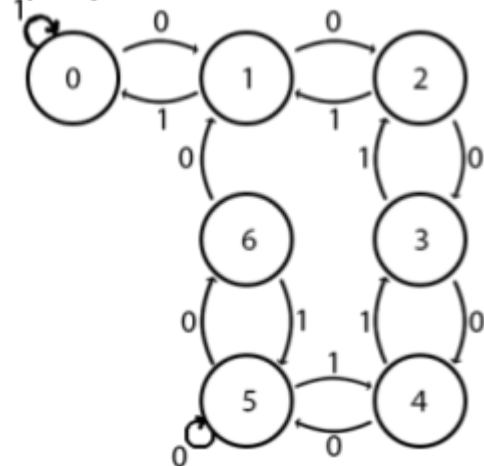
Figura 5: Movimento do agente no ambiente 1 por ação



Dependendo do estado inicial ele irá dirigir-se para o estado objetivo mais próximo (estado 0 ou estado 6), sendo que no estado 3 a decisão é aleatória.

3.3.2. Segundo Ambiente

Figura 6: Movimento do agente no ambiente 2 por ação



No segundo ambiente o agente também irá andar em direção ao estado objetivo (neste caso apenas o estado 0) sendo que ele poderia ficar parado no estado 5, contudo a recompensa (zero) não o motiva a escolher esta opção mais vezes. Na nossa analogia, o robot encontra uma estação de carregamento que não é dele então não é incentivado a utilizá-la.