



Grupo T1G18

Agentes e Inteligência Artificial Distribuída

Catarina Figueiredo	- up201606334
Lucas Vieira Stein	- up201606398
Pedro Tavares	- up201406991

DESCRIÇÃO

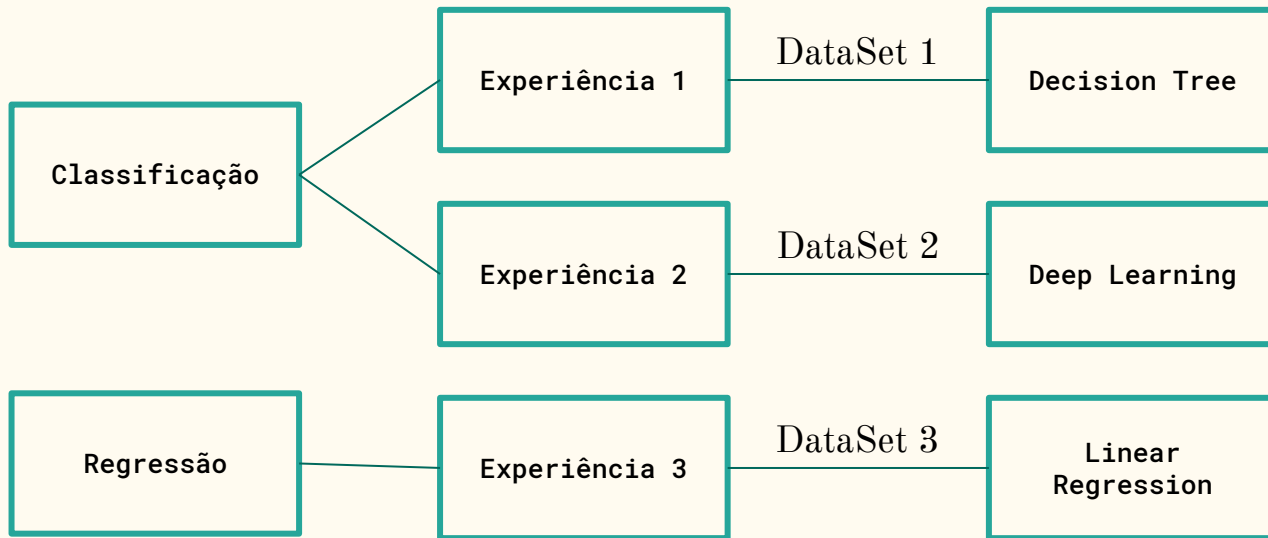
UberEats™ é um serviço de entrega de refeições ao domicílio. O cliente exerce um pedido a um restaurante, através de uma aplicação, esse pedido é depois atribuído a um funcionário da Uber que fica responsável pelo seu transporte.

Existe uma necessidade de avaliar o desempenho de cada restaurante e interpretar quais são os fatores mais relevantes que afetam a decisão do cliente. Estes fatores podem variar mediante a sua localização, o critério de escolha de cada cliente e a população local.

Problema 1 (Classificação): Como é que a localização de um restaurante, a população de clientes e restaurantes afetam o seu desempenho? Que impacto terá o critério de escolha de um cliente?

Problema 2 (Regressão): Que fatores são mais relevantes a determinar o tempo de entrega de cada condutor na sua deslocalização ao restaurante e ao domicílio?

Experiências Realizadas





Experiências 1 & 2

Apresentação

Catarina Figueiredo	- up201606334
Lucas Vieira Stein	- up201606398
Pedro Tavares	- up201406991

Experiências Realizadas

Experiência 1

Objetivo:

Problema de classificação (**Problema 1**)

Variáveis independentes:

nrClients: *int*, número de clientes global

nrRestaurants: *int*, número de restaurantes global

%Crit0,1,2: *real*, percentagem de clientes para cada critério.

clientsQ1,2,3,4: *int*, distribuição de clientes por quadrante.

restQ1,2,3,4: *int*, distribuição de restaurante por quadrante.

mediumRankRest: *int*, ranking do restaurante

mediumPrice: *int*, preço médio dos pratos do restaurante

QtoAnalyse: *int*, quadrante em que o restaurante está localizado.

Variável dependente:

RestRating: *string*, classificação do desempenho do restaurante (mau, médio, bom, muito bom)

Experiência 2

Objetivo:

Problema de classificação (**Problema 1**)

Variáveis independentes:

Mesmo que na experiência 1.

Variável dependente:

Mesmo que na experiência 2.

Nesta experiência optamos por usar um dataset com maior variabilidade nos resultados do **RestRating**.

Detalhes sobre o DataSet 1 & 2

Cada linha dos datasets recolhidos correspondem a um restaurante selecionado aleatoriamente. Este restaurante vai estar localizado em um dos quatro quadrantes do nosso mapa, declarado no atributo **QtoAnalyse**. Relacionados serão também os atributos **mediumRankRest** e **mediumPrice** que respectivamente correspondem à qualidade do restaurante e o preço médio da sua ementa.

O **RestRating** é a determinação do desempenho desse restaurante em questão. Vamos considerar que um restaurante que faz mais de 6 pedidos é “Muito Bom”, mais de 4 “Bom”, mais de 2 “Médio” e entre 0-1 “Mau”.

Processos do Rapidminer

Os processos utilizados no **rapidminer** para cada experiência foram os seguintes:

Experiência 1: Decision Tree

Experiência 2: Deep Learning

Experiência 3: Linear Regression

Estatísticas da Experiência 1

<div>Label</div> <div>RestRating</div>	Polynomial	0	Least Muito Bom (50)	Most Mau (487)	Values Mau (487), Medio (351), ...[2 more]			
<div>Prediction</div> <div>prediction(RestRating)</div>	Polynomial	0	Least Muito Bom (2)	Most Mau (875)	Values Mau (875), Medio (118), ...[2 more]			
<div>Confidence_Medio</div> <div>confidence(Medio)</div>	Real	0	<div></div> <div>Open visualizations</div>		Min 0	Max 1	Average 0.351	Deviation 0.093
<div>Confidence_Mau</div> <div>confidence(Mau)</div>	Real	0	<div></div> <div>Open visualizations</div>		Min 0	Max 1	Average 0.487	Deviation 0.192
<div>Confidence_Bom</div> <div>confidence(Bom)</div>	Real	0	<div></div> <div>Open visualizations</div>		Min 0	Max 1	Average 0.112	Deviation 0.097
<div>Confidence_Muito Bom</div> <div>confidence(Muito Bom)</div>	Real	0	<div></div> <div>Open visualizations</div>		Min 0	Max 1	Average 0.050	Deviation 0.068

Estatísticas da Experiência 2

<div>Label</div> <div>RestRating</div>	Polynomial	0	Least Muito Bom (53)	Most Medio (503)	Values Medio (503), Mau (321), ...[2 more]		
<div>Prediction</div> <div>prediction(RestRating)</div>	Polynomial	0	Least Bom (49)	Most Medio (500)	Values Medio (500), Mau (366), ...[2 more]		
<div>Confidence_Mau</div> <div>confidence(Mau)</div>	Real	0	<div></div> <div>Open visualizations</div>	Min 0.000	Max 0.984	Average 0.365	Deviation 0.294
<div>Confidence_Bom</div> <div>confidence(Bom)</div>	Real	0	<div></div> <div>Open visualizations</div>	Min 0.000	Max 0.688	Average 0.119	Deviation 0.130
<div>Confidence_Medio</div> <div>confidence(Medio)</div>	Real	0	<div></div> <div>Open visualizations</div>	Min 0.015	Max 0.930	Average 0.434	Deviation 0.218
<div>Confidence_Muito Bom</div> <div>confidence(Muito Bom)</div>	Real	0	<div></div> <div>Open visualizations</div>	Min 0.000	Max 0.927	Average 0.081	Deviation 0.172

Análise do Rapidminer

Decision Tree		Deep Learning	
EXPERIÊNCIA 1	Accuracy	Accuracy	EXPERIÊNCIA 2
	53.70%	63.40%	
EXPERIÊNCIA 1	Classification Error	Classification Error	EXPERIÊNCIA 2
	46.30%	36.60%	

Conclusões

Para a elaboração destas experiências o grupo gerou datasets com 1000 linhas de informação. Para uma obtenção de resultados mais fiáveis poderíamos ter atribuído uma amostra com maior número de entradas. Apesar de tudo, conseguimos já formular uma conclusão em relação aos dados obtidos.

Observamos que na análise **Deep Learning** conseguimos uma previsão muito mais fiável, com uma percentagem de erro mais reduzida.

Dada a simplicidade do nosso sistema de localização não tivemos muita variedade de atributos para a nossa análise. Adicionando mais complexidade teria de certa forma contribuído para resultados mais interessantes.

Experiência 3

Apresentação

Catarina Figueiredo	- up201606334
Lucas Vieira Stein	- up201606398
Pedro Tavares	- up201406991

Experiência Final

Experiência 3

Objetivo:

Problema de regressão (Problema 2)

Variáveis independentes:

nrClients: *int*, número de clientes global

nrRestaurants: *int*, número de restaurantes global

nrDrivers: *int*, número de condutores global

%Crit0,1,2: *real*, percentagem de clientes para cada critério.

clientsQ1,2,3,4: *int*, distribuição de clientes por quadrante.

restQ1,2,3,4: *int*, distribuição de restaurante por quadrante.

drivQ1,2,3,4: *int*, distribuição de condutores por quadrante.

Variável dependente:

AvgTime: *int*, tempo médio dos condutores de uma simulação.

Análise Rapidmine:

Usando o processo de Linear Regression

Correlation: **0.623**

Root Mean Squared Error: **11.744 +/- 0.0**

Root Relative Squared Error: **0.785** (Resultado está no intervalo entre 0 e 1 portanto é um bom modelo para esta previsão.)