INSPER – INSTITUTO DE ENSINO E PESQUISA

Ciência dos Dados

PROF. Fabio Ayres

Classificação e aplicação de naive bayes nos jogos da steam

Lucas Fukada

MARCELO MIGUEL

Eiki Yamashiro

Beatriz muniz

**SUMÁRIO**

SÃO PAULO, 19 de novembro de 2019

**1 INTRODUÇÃO/OBJETIVOS.......................................... 3**

**2 DATASET........................................................................... 3**

**3 PROCEDIMENTOS EXPERIMENTAIS........................ 4**

**4 RESULTADOS E DISCUSSÕES....................................... 6**

**5 CONCLUSÕES................................................................... 7**

**1 INTRODUÇÃO/OBJETIVOS**

O projeto consiste na utilização de Classificação e aplicação de Naive Bayes em jogos da plataforma Steam. Dessa forma, para a Classificação, o objetivo é de classificar o jogo em ser bom, muito bom ou ruim (utilizando a categoria de “Feedbacks positivos” e “Feedbacks negativos”), com base no preço do jogo, a quantidade de categorias que ele se encaixa, media de tempo jogado e a mediana de tempo jogado. Enquanto isso, para o Naive Bayes, o objetivo é o mesmo da Classificação, porem a intensidade do feedback dependeria apenas do seu nome.

**2 DataSet**

O DataSet foi retirado do site [kaggle](http://www.kaggle.com),

**3 PROCEDIMENTOS EXPERIMENTAIS**

Os procedimentos experimentais serão divididos em alguns tópicos com o intuito de uma melhor organização dos métodos utilizados para a obtenção de resultados e os acontecimentos ocorridos no decorrer da experimentação.

3.1 TEMPO DE MANTIMENTO DO TERMÔMETRO NOS FORNOS

Inicialmente, para analisar a temperatura no termômetro, foi utilizado 1min, no entanto, por uma tentativa aleatória, deixou-se 3min e notou-se que a temperatura medida foi muito diferente com o tempo, mas a partir desse tempo a medição não se alterava de forma considerável. Portanto, a partir dessa medição, voltou-se todas as mensurações utilizando o método de 3min do TLV nos fornos.

3.2 AS TEMPERATURAS APONTADAS NO EQUIPAMENTO E AS MOSTRADAS PELO TLV

No início do experimento, a dupla utilizou outro termômetro, este que, por conta de um procedimento realizado indevidamente, foi descartado, já que ocorreu um choque térmico no vidro e acabou quebrando o equipamento, apontando a importância da sequência das medidas de temperatura. Por conta disso, os dados obtidos pelo gráfico somente contém os dados do experimento em que o termômetro se manteve intacto, no qual utilizamos a estratégia correta.

Tabela 1 – Medidas apontadas pelo termômetro intacto

|  |  |
| --- | --- |
| **TEMPERATURA REAL [°C]** | **TEMPERATURA NO TERMÔMETRO [°C]** |
| -2,8 | -0,4 |
| 0 | 0,1 |
| 16,3 | 15,8 |
| 32 | 30 |
| 46 | 45 |
| 65,8 | 62,3 |
| 80 | 78 |
|  |  |

3.3 CALIBRAÇÃO E CÁLCULO DA EQUAÇÃO DE CALIBRAÇÃO DO TERMÔMETRO

3.3.1 CÁCULO PELO EXCEL

Utilizando os dados anteriores no Excel e gerando um gráfico de dispersão com os dados, seleciona-se a opção de linha de dispersão para ter uma linha de calibração, e como é apontado, a equação é polinomial de grau 1, do tipo , sendo os coeficientes a = 0,9517 e b = y = 0,7102.

Gráfico 1 – EXCEL: TEMP. MEDIDA X TEMP. REAL

3.3.2 CÁLCULO POR MMQ

Para confirmar a equação da linha de tendência, o grupo utilizou o Método dos Mínimos Quadrados (MMQ). Utilizando as fórmulas abordadas pela temática, obteve-se a função . Isto é, diante dessa tentativa, obteve-se um valor totalmente parecido, diferenciando-se apenas no 4° termo após a vírgula, demonstrando que a estratégia foi eficiente e resultou no que era esperado.

**4 RESULTADOS E DISCUSSÕES**

Os resultados obtidos fizeram com que os objetivos almejados no início do experimento fossem conquistados, e conseguiu-se calcular a equação de calibração do termômetro com o número de série 53169/15, sendo ela , alterando um pouco apenas no MMQ, no entanto sem atrapalhar completamente o que seria, já que mudou-se apenas a quarta casa decimal após a vírgula.

Ademais, após a experimentação, obteve-se o gráfico de dispersão no Excel (Gráfico 1) e o que será abordado abaixo como forma de comparação dos dados.

Gráfico 2 – Relação entre as temperaturas

Diante disso, pode-se observar de uma maneira divergente os mesmos dados obtidos anteriormente.

Além de tudo, foi obtida a temperatura do ambiente do laboratório após a calibração. Antes de calibrar, era de 23ºC; após, com a função pelo MMQ, a temperatura seria de 23,4°C.

**5 CONCLUSÕES**

Após a finalização e a análise experimental, conclui-se que o MMQ se torna muito relevante para uma calibração, equivalendo ao que se mostra no Excel, demonstrando que o experimento foi relevante e com os resultados obtidos, isto é, resolveu-se o que foi proposto e conseguiu-se, com êxito, calcular a temperatura do ambiente do laboratório e concluir o experimento.