**Conceitos Estatísticos para IA – Turma 24IA**

**Trabalho 5**

Habilidades desenvolvidas: Modelo de Regressão Linear Múltipla. Regressão Logística.

**Parte 1**

1. Relacione a Estatística com a Definição do conceito estatístico:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estatísticas |  |  | Definição |
| (a) | Y |  | ( ) | Intercepto do modelo de regressão (nível médio) |
| (b) | Β0 |  | ( ) | Variável predito do modelo |
| (c) | Multiple R2 |  | ( ) | Diferença entre a resposta e o valor estimado pelo modelo (Y-Y estimado) |
| (d) | Adjusted R2 |  | ( ) | Acurácia do modelo de regressão linear simples. |
| (e) | Β1 |  | ( ) | Acurácia do modelo de regressão linear múltipla. |
| (f) | Resíduo padronizado |  | ( ) | Regressão linear simples |
| (h) | Y estimado |  | ( ) | Regressão linear múltipla |
| (i) | Y=β0+β1x |  | ( ) | Valor estimado pelo modelo de regressão linear |
| (j) | Y=β0+β1x1+β2x2 |  | ( ) | Coeficiente de regressão |
| (k) | Resíduo |  | ( ) | Resíduo é transformado em nova variável com média igual 0 e desvio padrão igual 1 |
| (l) | X |  | ( ) | Variável resposta do modelo |
|  |  |  |  |  |

1. Use os dados *Bike\_Sharing.csv* para construir as análises descritivas, correlação, associação e modelo preditivo para previsão do número de bikes alugadas por mês.

Descrição:

Os sistemas de compartilhamento de bicicletas são uma nova geração de aluguel de bicicletas tradicional, onde todo o processo de associação, locação e devolução tornou-se automático. Através destes sistemas, o usuário pode facilmente alugar uma bicicleta a partir de uma determinada posição e retornar em outra posição. Atualmente, existem cerca de 500 programas de compartilhamento de bicicletas em todo o mundo, compostos por mais de 500 mil bicicletas. Hoje, existe um grande interesse nesses sistemas devido ao seu importante papel no trânsito, questões ambientais e de saúde.

Fonte de dados: <https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset>

Classifique o tipo de variável

|  |  |  |  |
| --- | --- | --- | --- |
| **Variável** | **Descrição** | **Tipo de Variável** | **Regra de data mining (Preditora ou Target)** |
| **instant** | record index | ID | - |
| **dteday** | date | ID | - |
| **season** | season (1:winter, 2:spring, 3:summer, 4:fall) |  |  |
| **yr** | year (0: 2011, 1:2012) |  |  |
| **mnth** | month ( 1 to 12) |  |  |
| **holiday** | weather day is holiday or not |  |  |
| **weekday** | day of the week |  |  |
| **workingday** | if day is neither weekend nor holiday is 1, otherwise is 0. |  |  |
| **weathersit** | 1: Clear, Few clouds, Partly cloudy, Partly cloudy; 2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist;3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds;4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog |  |  |
| **temp** | Normalized temperature in Celsius. The values are derived via (t-t\_min)/(t\_max-t\_min), t\_min=-8, t\_max=+39 (only in hourly scale) |  |  |
| **atemp** | Normalized feeling temperature in Celsius. The values are derived via (t-t\_min)/(t\_max-t\_min), t\_min=-16, t\_max=+50 (only in hourly scale) |  |  |
| **hum** | Normalized humidity. The values are divided to 100 (max) |  |  |
| **windspeed** | Normalized wind speed. The values are divided to 67 (max) |  |  |
| **casual** | count of casual users |  |  |
| **registered** | count of registered users |  |  |
| **cnt** | count of total rental bikes including both casual and registered |  |  |

1. Definição do problema:
2. Objetivo do estudo:
3. Apresente as tabelas descritivas e gráficos das variáveis. Interprete os resultados.
4. Quais variáveis foram normalizadas? Por quê? Apresente a fórmula utilizada.
5. Apresente a análise de associação e correlação de Pearson. Quais variáveis têm correlação com a variável resposta?
6. Construir o modelo de regressão linear múltipla.
7. Qual a acurácia do modelo?
8. Construir o simulador do modelo ajustado no item 6).

**Data de entrega: 31/05/2022**

**Formato: Word**

**Regina Bernal**

**21/05/2022**