

INE5649 – Técnicas Estatísticas de Predição
Prof. Luiz Ricardo Nakamura
Contato: luiz.nakamura@ufsc.br

TRABALHO EM GRUPO

MODELO DE REGRESSÃO LINEAR MÚLTIPLA

RESUMO EXPANDIDO: 10 PONTOS VIDEO: 10 PONTOS

Este trabalho em grupo está previsto em nosso plano de ensino e irá compor 60% da média final desta disciplina. Para sua realização, siga integralmente todas as instruções contidas neste arquivo.

Instruções

1. O trabalho deve ser realizado, necessariamente, em grupos de 2 a 4 estudantes e será composto pela entrega de um **resumo expandido** e pela confecção de um **vídeo** que será socializado com a turma.
2. O **resumo expandido** trata-se de um documento escrito, utilizando fonte *Arial*, tamanho 11, espaçamento 1,5 entre linhas, margens 2,0 cm (superior, inferior, esquerda e direita). O resumo deverá ser composto, necessariamente, pelas seguintes seções:
 - (a) Título;
 - (b) Identificação dos membros do grupo;
 - (c) Introdução, justificativa e objetivos;
 - (d) Materiais e métodos;
 - (e) Resultados e discussões;
 - (f) Considerações finais;
 - (g) Referências bibliográficas (formato ABNT).
3. O **resumo expandido** pode conter, no máximo, 10 páginas. O arquivo com o trabalho deve ser enviado em formato *.pdf diretamente no sistema Moodle, observando a data limite estipulada (apenas UM integrante deve postar o arquivo no sistema).
4. O **vídeo** deve conter, no máximo, cinco minutos e deve sintetizar o resumo expandido, apresentando os principais resultados e suas discussões. Deve ser disponibilizado, por exemplo, no YouTube como vídeo não listado, de modo que todos da turma possam ter acesso à apresentação. A URL do vídeo deve ser disponibilizada no momento de submissão do resumo expandido.
5. O trabalho deve ser entregue na data estipulada. Entregas recebidas com atraso (de até 24 horas) terão valor máximo de 8,0 pontos. Após as 24 horas de atraso, não serão aceitas mais entregas.

6. Os seguintes itens devem ser desenvolvidos:

- (a) Apresentar um problema real (escolha um tema de sua área, ou de sua preferência) que pode ser analisado via análise de regressão linear múltipla.
- (b) Defina claramente e descreva sobre a variável dependente e as variáveis independentes (pelo menos três, sendo pelo menos uma quantitativa). Justifique o porquê dessas variáveis (como elas poderão responder ao problema descrito na alínea a). É requerido um número mínimo de 30 unidades observacionais (linhas) no conjunto de dados, que pode ser extraído de uma alguma fonte que deve, necessariamente, ser citada.
- (c) Realize uma análise descritiva e exploratória (construção de gráficos adequados e cálculo de medidas de resumo) para a variável resposta em estudo.
- (d) Apresente as medidas de resumo para as variáveis explicativas escolhidas.
- (e) Faça uma análise com diagramas de dispersão (ou gráficos de caixa para variáveis categóricas) entre as variáveis dependente e cada uma das variáveis independentes em estudo:
 - Todas as relações com a variável dependente parecem lineares? Se sim, os coeficientes de correlação linear indicam que essa relação é forte?
 - Há pontos discrepantes? Os gráficos sugerem alguma transformação?
 - Faça as devidas mudanças sugeridas a partir dos diagramas para a sequência do trabalho. Se fizer alguma transformação ou retirada de dados discrepantes, explicita no texto e refaça os diagramas.
- (f) Obter a equação de regressão e outras estatísticas de interesse usando um software apropriado. Fazer uma análise crítica sobre o modelo ajustado:
 - O modelo é útil? (Teste F de utilidade do modelo) (utilizar $\alpha = 5\%$).
 - Todas as variáveis independentes têm efeitos significativos (utilizar $\alpha = 5\%$) (Teste t sobre os coeficientes).
 - A análise dos resíduos (gráficos) sugere que o modelo está adequado?
- (g) Faça os devidos ajustes no modelo, se necessário.