**Lista de Regressão Logística – Análise de Dados 4  
Marcos Rodrigues de Oliveira Júnior**

**1-A)**

As variáveis significativas, de acordo com o p-valor (menor que 0,05) são: t\_emprego, t\_endereco e divida\_cc

Pelo método dos coeficientes, mais longes de 0 são: divida\_cc, t\_emprego e divida

Pelo método do KBest, usando um seletor, todas são importantes.

Escolheremos interpretar pelo p-valor

**C)**

As variáveis escolhidas continuam sendo significantes.

Temos que o modelo, diminuiu sua acurácia, antes (com todas as variáveis) estava com 80% depois do ajuste prático 84,5%, agora está com 78,3% e depois do ajuste prático 82%.

**Pseudo R-quadrado**: **0.225**. Isso indica que o modelo explica aproximadamente 22,5% da variação da variável dependente.

**Log-Likelihood: -131.57.** Representa a probabilidade do modelo de ajustar os dados. Valores mais negativos indicam um ajuste menos bom.

**LLR p-value**: **2.5171e-17**. Um valor muito pequeno indica que o modelo como um todo é estatisticamente significativo. Ou seja, as variáveis incluídas no modelo têm impacto significativo na previsão.

Com uma sensibilidade de: 36%

Com uma especificidade de: 99%

Gráfico, Gráfico de linhas

O conteúdo gerado por IA pode estar incorreto.

**D)**

De acordo com o AUC de 0,672. Temos que o modelo não é ideal, pois quanto mais alto melhor, e sobre a curva, esperamos que quando mais próxima do canto superior esquerdo, melhor, podemos ver que ela não faz isso.

**E)**

Utilizando o intercepto fornecido pelo modelo como 0.12094549 e os coeficientes obtidos na regressão logística, podemos calcular a probabilidade de default de um indivíduo com as seguintes características: 3 anos de tempo de emprego, 5 anos de residência no mesmo endereço, e uma dívida no cartão de crédito de $0.70 mil. Para realizar esse cálculo, utilizamos a fórmula da regressão logística, chegando no resultado de **68.82%**.