

### 1. Case RH - Retenção de funcionários

Estamos falando que Big Data usada de modo científico e metodológico não só pode nos permitir entender o que os dados estão nos dizendo sobre o passado, mas também podemos utilizá-los para prever o futuro. Os grupos particulares de funcionários em risco de sair e aqueles grupos de talentos que serão particularmente difíceis de reter ou adquirir no futuro.

Baseado em alguns itens, facilmente disponibilizados por departamentos de RH, modelos algorítmicos são capazes de extrair análises profundas que vão permitir empresas decidirem sobre a tomada de ações corretas para imediatamente melhorar as taxas de retenção do seus funcionários.

Imagine que você está em um projeto juntamente com a equipe de RH com o objetivo de melhorar as taxas de retenção de sua empresa. Crie um modelo que represente o Job Satisfaction de um profissional de TI a partir de características da companhia e outros dados obtidos pela pesquisa. Utilize os dados da pesquisa de 2018 do Stackoverflow.

Apresente a intuição que suporta o modelo escolhido. Comprove a assertividade deste modelo através de experimentos.

Recrutamento é fundamental. Datasets de recrutamentos passados combinados com d

- detecção de talentos em potencial,
- aumento da taxa de sucesso em recrutamento

- Aplicar Análise preditiva eficientemente em dados de RH

- prever ociosidade e acidentes de trabalho,

- Big data analytics em RH permite o entendimento de forma acurada das motivações

Sabendo-se que você trabalha em cooperação com a equipe de marketing de uma startup que promete melhorar as taxas de recrutamento para empresas de tecnologia através

Crie um modelo que represente o Salary de um profissional de TI a partir de características da empresa e outras características desejáveis dos funcionários. Utilize os dados disponibilizados na pesquisa de 2018 do Stackoverflow.

Apresente a intuição que suporta o modelo escolhido. Comprove a assertividade deste modelo através de experimentos.

1. Carroq

- remover linhas com missing
- codificar variáveis categóricas como fatores

Link da survey: <https://insights.stackoverflow.com/survey>

Link da base semi-tratada: [https://www.dropbox.com/s/vy0gyv6bs9dn63f/dados\\_s](https://www.dropbox.com/s/vy0gyv6bs9dn63f/dados_s)

OBS: Vocês deverão utilizar a base semi-tratada para o proj