

Zakres tematyczny szkolenia.

Szkolenie (w wymiarze minimum 240 godzin z trenerem) będzie przygotowywało uczestników do podjęcia pracy na stanowisku Junior Data Scientist i będzie swoim zakresem obejmować minimum następujące zagadnienia:

1. Narzędzia pracy:
 - PyCharm/VSC, Jupyter Notebook, GitHub
 - Git
 - Tableau

2. SQL i bazy danych
 - bazy danych (projektowanie, typy danych, DML, tworzenie tabel)
 - podstawy programowania SQL
 - składnia języka SQL
 - przetwarzanie danych w SQL

3. Python
 - podstawy Pythona
 - funkcje
 - konwersje
 - typy i struktury danych
 - zmienne
 - biblioteki
 - instrukcje warunkowe
 - moduły i pakiety
 - klasy
 - wyjątki
 - wyrażenia regularne
 - środowisko wirtualne

- logowanie
 - operacje na plikach
 - przetwarzanie danych w Pythonie
 - czyszczenie danych
 - analiza danych
4. Wprowadzenie teoretyczne do statystyki
- mediana, percentyle, kwartyle
 - information value
 - predykcje
5. Machine Learning
- wstęp do Machine Learning
 - podział danych na treningowe i testowe
 - klasyfikacja, detekcja i segmentacja
 - regresja liniowa, logistyczna i klasteryzacja
 - Scikit-learn, Pandas, numpy, OpenCV, matplotlib
 - mierzenie skuteczności modeli
 - modele ML:
 - drzewo decyzyjne i random forest
 - KNN (K-Nearest Neighbors) i K-means
 - SVM (Support Vector Machine)
 - XGBoost
 - porównanie modeli
6. Deep Learning
- wstęp do Deep Learning
 - tworzenie sztucznych sieci neuronowych:
 - konwolucyjne sieci neuronowe
 - rekurencyjne sieci neuronowe
 - praca z modelami DL:
 - funkcje aktywacji i kosztu
 - zanikanie gradientu
 - oszukiwanie sieci

- wizualizacja sieci
 - uczenie nienadzorowane
 - uczenie ze wzmocnieniami
- TensorFlow