

# RTBHOUSE =

## Data Scientist

### Team Model

#### TWOJE ZADANIA:

- Projektowanie i programowanie modeli, najczęściej głębokich sieci neuronowych, służących do przewidywania zachowań i preferencji internautów
- Dostosowywanie budowanych rozwiązań do wymagań produkcyjnego systemu (m.in optymalizacja latencji czy zużycia zasobów)
- Implementacja, dostosowywanie i usprawnianie wyników prezentowanych w najnowszych pracach naukowych
- Przeprowadzanie testów A/B nowych rozwiązań (jeśli Twój model jest lepszy, to zostaje na produkcji)
- Ciągła optymalizacja istniejących narzędzi i procesów
- Odnalezienie dla siebie takiego miejsca pomiędzy pozycją Research Scientist a Big Data Engineer, które najefektywniej wykorzystuje Twoje umiejętności

#### PRZYKŁADOWE TEMATY:

- Wymiana modelu Field-aware Factorization Machines na model oparty o głębokie sieci neuronowe z lepszym wsparciem zmiennych ciągłych o różnorodnych rozkładach
- Analiza znaczenia niewykorzystywanych dotychczas informacji
- Eksperymenty z wykorzystaniem transfer learningu
- Zupełnie nowe podejście do generowania rekomendacji produktowych, które jest wolne od proxy targetów i wykorzystuje contrastive loss function
- Opracowanie metody optymalizacji wspierającej problemy regresyjne o różnorodnych rozkładach
- Rozbudowa modelu o nowe dane wejściowe
- Stabilizacja procesu uczenia produkcyjnego (obsługa tymczasowej niedostępności sieci, checkpointy i automatyczne wznowianie uczenia)
- Niskopoziomowa optymalizacja procesu uczenia i ewaluacji modeli (fast NVMe sampler for DNN training, CUDA FFM training + AVX FFM inference)
- Automatyczna weryfikacja poprawności i kompletności danych (wykrywanie anomalii itp.)

#### WYKORZYSTYWANE TECHNOLOGIE:

- **PyTorch, Torch**
- **Nvidia CUDA**
- **Python, Pandas, Numpy, Jupyter**
- **BigQuery, Hadoop, Crunch, Spark**
- **Java, Scala**

**NASZE WYMAGANIA:**

- Zainteresowanie tematyką uczenia maszynowego, szczególnie Deep Learningiem
- Biegłość w programowaniu
- Znajomość statystyki i rachunku prawdopodobieństwa
- Umiejętność dobierania najlepszych środków do osiągnięcia założonych celów, np. poprzez elastyczne stosowanie języków programowania, frameworków, narzędzi itp.
- Chęć do eksperymentowania i wyciągania wniosków

**DODATKOWYMI ATUTAMI BĘDĄ**

- Bardzo dobra znajomość zagadnień związanych z sieciami neuronowymi i Deep Learningiem od strony teoretycznej i wykorzystania w praktyce
- Praktyczna znajomość wybranych frameworków takich jak Torch, PyTorch, czy TensorFlow
- Doświadczenie w programowaniu z użyciem GPU
- Znajomość wybranych technologii Big Data takich jak: Hadoop, Crunch, Spark czy Google Big Query
- Znajomość podstawowych zagadnień (typowych problemów i metod ich rozwiązywania) z obszarów: high availability, high performance i high scalability

**W ZAMIAN OFERUJEMY:**

- Bardzo atrakcyjne wynagrodzenie
- Pracę w zespole pasjonatów posiadającym doświadczenie w uczeniu maszynowym, którzy z chęcią dzielą się swoją wiedzą i umiejętnościami
- Wyjątkowo elastyczne warunki pracy – pracujesz kiedy Ci wygodnie i poświęcasz tyle czasu, ile możesz; w dużej części możesz pracować zdalnie
- Sprzęt i oprogramowanie, jakiego potrzebujesz
- Możliwość publikacji swoich wyników
- Dostęp do najnowocześniejszych technologii i możliwość realnego wykorzystywania ich w projekcie o dużej skali i wysokiej dynamice
- Możliwość wykorzystania posiadanej wiedzy i kompetencji w praktycznych zastosowaniach – przy optymalizacji algorytmów obsługujących setki milionów internautów i kupujących miesięcznie miliardy odsłon reklamowych w modelu RTB, bazując na obszernych zbiorach danych
- Efekty Twojej pracy widoczne od razu w wynikach biznesowych firmy

Więcej informacji: [techblog.rtbhouse.com/jobs/](http://techblog.rtbhouse.com/jobs/)

Kontakt: [it-jobs@rtbhouse.com](mailto:it-jobs@rtbhouse.com)