

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/

Kontakt: it-jobs@rtbhouse.com