Talgat Omarov   Data Scientist, M.Sc.

Edmonton, Canada | +1 (587) 594-6537 | contact@talgatomarov.com | GitHub | LinkedIn

SKILLS

- **Prog. Languages**: Python, PyTorch, Scikit-Learn, Pandas, Numpy, R, tidyverse, mgcv, fable, rstan.
- **Database Systems**: SQL, MySQL, MongoDB, DuckDB.
- **Skills:** Statistical Modelling, Machine Learning, Natural Language Processing, Cloud Computing (AWS), Data Visualization (Tableau, Quarto), Time Series, Regression Analysis, Bayesian Modeling, ETL.

EXPERIENCE

Data Scientist   **May 2023 - Present**
EngageFirst Management Consultants   Edmonton, Canada
Advanced Analytics Implementation for Alberta's Ministry of Children and Family Services (CFS)

- **Stakeholder Engagement**: Collaborated with CFS stakeholders to gather requirements, aligning with organizational goals, and communicated analytical findings clearly to support informed decision-making.
- **Predictive Modeling**: Developed time-series forecasting models (ARIMA, ETS) in R to predict children and youth facility placements across regions and types over a three-year horizon for strategic planning.
- **Risk Analysis**: Utilized Cox proportional hazards model in R to identify significant predictors of children and youth facility placements, improving the understanding of historical trends in administrative data.
- **Secure AWS Environment**: Established a secure and scalable AWS infrastructure for data storage and processing, ensuring confidentiality and meeting the high computational demands of the project.
- **Data Quality Assessment**: Led a comprehensive assessment of data accuracy, completeness, and consistency, delivering actionable recommendations to enhance data integrity and overall quality.

Graduate Research and Teaching Assistant   **Sep 2021 - April 2023**
University of Alberta   *Edmonton, Canada*

- **Natural Language Processing System**: Developed an advanced multilingual idiomaticity detection system using PyTorch and Transformers, achieving 3rd place out of 16 teams in the SemEval 2022 competition.
- **Research Publication**: Published a peer-reviewed paper on the system in the workshop proceedings.
- **Lab Instruction**: Led interactive lab sessions for 50+ students on data structures, database design, relational algebra, SQL and MongoDB, effectively enhancing their understanding of database concepts.

Data Scientist Intern   **Jun 2021 - Aug 2021**
Kaspi Bank   *Astana, Kazakhstan*

- **Recommendation Model**: Developed a prototype recommendation model for a leading e-commerce platform in Kazakhstan using Python and matrix factorization techniques to improve customer experience.
- **Data Science Bootcamp**: Graduated from Kaspi Lab Data Science bootcamp, specializing in Python, data analysis, and machine learning, with a focus on data-driven decision-making.

Data Engineer Intern   **May 2019 - Jul 2019**
KazData   *Astana, Kazakhstan*

- **ETL Pipeline Development**: Developed ETL pipelines using Python and SQL to extract, transform, and load data from multiple sources, including MS Access, MS Excel, CSV, and MySQL, into the data warehouse.
- **Data Management**: Managed over 8 million high-quality data entries, which were successfully sold to clients for strategic decision-making purposes.
- **Data Visualization:** Implemented a dynamic data visualization module using PHP and Highcharts.js, transforming complex datasets into intuitive visuals to boost client engagement and satisfaction.

EDUCATION

University of Alberta   **Sep 2021 - May 2023**
M.Sc. in Computing Science   Edmonton, Canada

- Cumulative GPA: **3.92/4.00.** Graduate Research Assistant Fellowship.

Nazarbayev University   **Aug 2017 - Jun 2021**
B.Sc. in Computer Science   Astana, Kazakhstan

- Cumulative GPA: **3.81/4.00**. Full scholarship.