

# ALI SHALBAYEV

Nur-Sultan, Kazakhstan

☎ +7 777 977 7242

✉ [alishalbayev1@gmail.com](mailto:alishalbayev1@gmail.com)

🌐 [linkedin.com/in/alishalbayev](https://www.linkedin.com/in/alishalbayev)

🐙 [github.com/alishalbayev](https://github.com/alishalbayev)

## Work Experience

### Data Analyst/Machine Learning Engineer

April 2022 – Present

*Presidential Administration of Republic of Kazakhstan*

*Nur-Sultan, Kazakhstan*

- Built a dashboard describing the performance of government agencies.
- Used NLP algorithms to preprocess text in both Kazakh and Russian to cluster similar sentences or texts.
- Identified risks and wrote algorithms for identifying suspicious purchases in public procurement.

Used technologies: *PowerBI, GraphQL, FAISS, LaBSE, Tesseract-OCR, RuBERT.*

### Data Analyst

December 2021 – April

*Transtelecom JSC*

*Nur-Sultan, Kazakhstan*

- Created a market penetration dashboard where a total number of active clients per residential building is available, also each building index is shown with specific color which depends on its value.
- Made dashboard that shows the utilization of cloud resources in percentages and colors.
- Wrote an app with GUI in Python Tkinter module that receives excel files and converts them into text files with some structure changes.

Used technologies: *QlikSense, Python, SQL.*

## Projects

### Brain Tumor Classifier

August 2021 – Present

*Astana Hub Incubation*

*Nur-Sultan, Kazakhstan*

- Uploaded and splitted tumor images for train and test stages.
- Implemented preprocessing techniques.
- Built machine learning models for tumour classification.

Used technologies: *OpenCV, Matplotlib, SVC and LogisticRegression.*

### Apartment Price Predictor, Classification and Clustering

January 2021 – June 2021

*Related Course*

*Nur-Sultan, Kazakhstan*

- Identified and handled null values and highly correlated columns.
- Implemented feature selector methods for keeping only those columns that have good predictive qualities.
- Built machine learning models for Classification, Clustering and Linear Regression tasks. Implemented Hyperparameter Tuning.

Used technologies: *FancyImputer, RandomForestRegressor, GridSearchCV, LassoRegressor, RidgeRegressor.*

### Korkem-Telecom LLC

January 2021

*Data Science Solved*

*Nur-Sultan, Kazakhstan*

- Identified regular buses, taxis, parents who take children to high school using clustering model.
- Implemented feature extraction by grouping transport based on different features.
- Visualised on the map all sergeks in Nur-Sultan.

Used technologies: *KMeans, Folium, Geocoder, Geopy, Ipywidgets, Pandas.*

### Entropy Encoding App

January 2021

*Related Course*

*Nur-Sultan, Kazakhstan*

- Wrote an algorithm of Shannon-Fano and Huffman encoding for the lossless data compression.
- Implemented Hamming coding for error detection after the compression.
- Built a web application for the better UI/UX.

Used technologies: *Streamlit, Numpy, Pandas.*

## Education

### Astana IT University

Sep. 2019 – May 2022

*Bachelor of Science in Computer Science*

*Nur-Sultan, Kazakhstan*

### Nazarbayev Intellectual school of Physics and Mathematics

Sep. 2013 – June 2019

*Middle and High School*

*Taraz, Kazakhstan*