

ELENA DUBOVA

✉ elena@dubova.net
🌐 dubova.net
☎ (551) 208-1562
📍 New York, NY 10019
📄 /in/elenadubova
🔄 elena-du

A seasoned business development professional with rigorous training in data science who is passionate about discovering impactful business insights that accelerate business growth and ethical data use.

Tools

DATABASES

Relational
Document-oriented

LANGUAGES

R
Python
SQL

DATA ANALYSIS

Machine Learning
Deep Learning
Predictive and Inferential Analytics
Business Intelligence
Natural Language Processing
Web Scraping
Cloud Computing

TOOLS AND APPLICATIONS

Shiny
Tableau
Flask
Dash
Git
ER Studio

Certifications

MSCA: SQL 2016 Database Developer, Microsoft

Big Data with Hadoop and Spark, CloudxLab

Deep Learning Specialization, deeplearning.ai

AI: Implications for Business Strategy, MIT Sloan

Business and Financial Acumen, INSEAD

Education

Columbia University of the City of New York
Master of Science Applied Analytics 2019

Ivanovo State University
Master of Arts Economics 2010

Ivanovo State University
Master of Arts International Relations 2007

Experience

Columbia University of the City of New York New York, New York
Associate Adjunct Faculty Sept. 2019 to Current
Advise students on programming in R and statistical learning

Metis New York, NY
Data Scientist Mar. 2020 to June 2020

- Completed an immersive 12-week accredited data science bootcamp. Designed five projects in python from ideation to presentation of data and findings.
- Honed skills in programming, predictive modeling, natural language processing and machine learning. **See project section.**

Microsoft Munich, Germany
Go-To-Market Strategy Lead Sept. 2015 to Dec. 2016

- Transformed the ecosystem of Independent Software Vendors (ISV) to grow Microsoft Azure consumed revenue in 33 countries of Central and Eastern Europe.
- Executed co-marketing and co-sell motions through market research, cross-functional collaboration, and Go-To-Market programs

Sales Team Lead Moscow, Russia
Jan. 2012 to Aug. 2015

- Managed a team of 10+ sales executives and \$20M+ annual sales quotas in public sector and consulting services industries.
- Oversaw key C-level relationships, advised on closing deals over 1M, coached sales team, and ensured top performance of the team in terms of revenue and KPIs. Maintained robust rhythm of business and ensured accurate revenue forecasts of entrusted business.

Account Executive Moscow, Russia
July 2008 to Jan. 2012

- Served as a Microsoft sales executive to key enterprise customers in professional services, retail and transportation industries.
- Member of the High Potential and Key Talent programs that provide advanced training for top 1% and top 5% of employees respectively.

Business Analyst Moscow, Russia
Aug. 2007 to June 2008

- Managed sales pipeline and delivered forecast insights.
- Used customer segmentation methodologies to optimize marketing effectiveness.

Projects

Blogger Boost Application

Developed a Dash application with three functional blocks: a 3D community visualizer based on NMF topic modeling analysis, a smart dictionary operating on community-specific word embeddings, and a self-awareness dashboard that analyses emotional blueprint of a blog over time and predicts emotional intensity of future blogs.

TED Talk Recommender System

Combined topic modeling of the contents of TED Talks with the sentiment analysis of the audience response to provide most satisfying recommendations to users based on plain text input. Constructed an interactive web application to provide recommendations.

Prediction of Hostility Level of Military Interstate Dispute

Studied applicability of quantitative methods to political science. Compared performance of variety of classification algorithms to predict the level of hostility of the conflict given economic and historical data. Analysed different approaches to handling class imbalance in the classification problem.

Prediction of Views on Youtube Yoga Channels

Created scripts to obtain YouTube Data. Inferred dependencies between features as well as interaction effects in order to describe the formula for YouTube view counts for yoga videos. Explored the limits of linear regression in predicting number of views.

Optimizing Location for New Residential Construction

Used statistical analysis to predict an optimal construction site in New York City.

Algorithmic Image Quality Improvement, XGBoost and CNN

Created a CNN model using Python that performs resolution enrichment in terms of peak signal-to-noise ratio (PSNR).

Ad Listings Topic Modeling, LDA, and Sentiment Analysis

Investigated the most and the least successful Airbnb listings from a text analysis perspective using sentiment analysis and topic modeling. Differentiated between successful and unsuccessful listings using a composite indicator built from rating score, popularity measures and super host indicator.

Optical Character Recognition Software (OCR) Text Error Detection and Correction

Created an OCR post-processing procedure to enhance Tesseract OCR output. Used Rule-based techniques to identify incorrectly processed words; used topic modeling to perform context-sensitive error correction.

University Hacker Application with Shiny

Designed an application to help potential college applicants choose a university in the United States based on their interest and aspirations. Independently set up the structure for the whole app

Understanding Happiness with Text Mining

Investigated a correlation between the length of a happy moment description and its properties. Performed exploratory data analysis of demographic data using the length of happiness descriptions as a variable in question; applied text mining techniques to unveil topics by loquacious and taciturn individuals

Economic Impact of Smoking - Tableau Application

Compiled an analytical visualization to demonstrate the impact of smoking on the cost of healthcare and solutions individuals may consider

Detecting Abnormalities in ECG Signal

Capstone project at Columbia University; developed a deep learning model architecture aimed to detect arrhythmia in electrocardiogram (ECG) data