SZILVESZTER TÓTH

CONTACT

+36/70 933-0910

totszilveszter@gmail.com

Budapest

LinkedIn: www.linkedin.com/in/szilveszter-tóth-7b9862181

SKILLS

- Teamwork
- Communication
- Critical Thinking
- Computer Skills
- Organisational / managerial skills
- Job-related skills

LANGUAGES

- Mother language
  - Hungarian
- English (Fluent)
  - Listening (B2)
  - Reading (B2)
  - Speaking (B2)
  - Writing (B2)

OTHER SKILLS

Driving licence

B

PROFILE

I enjoy working with machine learning and AI algorithms to develop solutions that automate processes and improve efficiency. I focus on providing practical, profitable solutions and simplifying tasks through IT. With experience in handling and processing large datasets, I aim to extract valuable insights and use data more effectively to support business goals.

EDUCATION

Software Engineer Master's Degree 2017-2019

Budapest University of Technology and Economoics

Thesis:

Development of a general-purpose framework to support performing intustrial processes using mixed reality techniques

Software Engineer Bachelor's Degree 2013-2017

Budapest University of Technology and Economoics

Thesis:

Vehicle detection and tracking

SKILLS

Communication skills

- Good cooperation and communication skills which were gained from my previous studies and projects in the frame of daily communication with my colleagues and customers

Organisational / managerial skills; Job-related skills

- Participated in coordination of juniour colleagues
- Participated in leading and elaboration of DevOps procedures frequently

Computer skills

- Good command of Microsoft Office™ tools
- Programming Languages: Java, C, C++, Python, C#, Typescript, Javascript, SQL, Power Query, DAX, CUDA, Terraform
- Technologies: Azure Batch Service, Azure Functions, .NET, Docker, Linux, Bash, Unity, OpenCV, Svelte, FastAPI, Opentelemetry, Github Actions, Gitlab CI, Kubernetes, PowerBI, PostgreSQL, Google Cloud services (Cloud Workflow, Cloud Scheduler,Cloud Run, Bigquery), Azure services, AWS services(AWS lambda, AWS Glue, AWS Athena), Terraform, dbt
- Machine Learning Tools: NVIDIA TensorRT, NVIDIA Digits, Caffe, TensorFlow, Jupyter Notebook, Scikit-Learn, Numpy, Azure ML Studio, PyTorch, Databricks
- Exams: AWS Certified Cloud Practitioner

SZILVESZTER TÓTH

+36/70 933-0910

totszilveszter@gmail.com

WORK EXPERIENCE

Data Engineer

Szallas Group Zrt., Budapest, Hungary, https://szallas.group/ 2025 MAY - PRESENT

Jira data request ticket helper

- Developed a RAG-based automation tool for Jira, utilizing Large Language Models to resolve frequent SQL data requests autonomously.
- Enhanced LLM accuracy by injecting DDL (Data Definition Language) schemas into the RAG context, ensuring syntactically correct query generation.
- Streamlined ticket operations by automating the end-to-end Jira workflow, including ticket ingestion, processing, and resolution uploading.

Data Warehouse implementation in GCP

- Architected and deployed a scalable DWH on BigQuery, utilizing Terraform for 100% Infrastructure-as-Code (IaC) management.
- Engineered high-fidelity ingestion pipelines using Sequin for Change Data Capture (CDC) and Airbyte for snapshot-based synchronization from PostgreSQL.
- Orchestrated containerized ETL workloads on a self-managed GKE cluster, leveraging Argo Workflows for complex dependency management.
- Optimized transformation layers using dbt to implement modular Data Marts and Star Schema architectures.
- Automated CI/CD lifecycles via GitLab and Artifact Registry to streamline Docker image deployments and dbt model updates.
- Established Data Governance frameworks by integrating Google Cloud Dataplex with automated Metadata tagging via Terraform.

GCP Data Platform – Implentation of external data ingestions in GCP

- Designed ingestion workflows with Cloud Run Jobs and Cloud Scheduler.
- Automated ETL deployments using GitLab CI/CD and Artifact Registry.
- Provisioned and managed infrastructure entirely with Terraform.
- Built DBT-based data pipelines with dependency-aware models on BigQuery.

SZILVESZTER TÓTH

+36/70 933-0910

totszilveszter@gmail.com

WORK EXPERIENCE

Data Scientist

Zenitech Ltd., Budapest, Hungary, www.zenitech.co.uk 2019 FEBRUARY - 2025 APRIL

PoC – UWB-Based Localization for Subway Systems

- Selected DW3000-based UWB modules suited for subway environments.
- Implemented SS-TWR algorithm on embedded devices to measure distances accurately.
- Applied least squares multilateration to compute initiator's 2D position.
- Built a live display interface to show anchor and initiator positions on a subway map.
- Validated system performance in conditions with multipath and partial line-of-sight.

The Hut Group – MLOps tools

- Developed and optimized MLOps tools to accelerate the deployment of data scientists' models to production environments.
- Implemented a remote execution framework for efficient GPU utilization on Google Cloud, enabling remote invocation of functions.
- Authored Request for Comments (RFCs) to clarify the architecture of tools for customers, enhancing their understanding.
- Designed and implemented a REST-based machine learning service framework to streamline the deployment of models created by data scientists.
- Created a machine learning service abstraction layer using Open Inference Protocol v2, integrated OpenTelemetry for observability, and developed an intuitive interface for service creation.
- Automated deployment of machine learning services to Google Kubernetes Engine (GKE) using GitHub Actions for Continuous Integration (CI).
- Built metrics dashboards with PromQL to monitor and visualize service performance.
- Integrated machine learning services with Triton Inference Server to support efficient model serving.

Graphisoft – Support agent

- Making human resource use more efficient with the development of a support agent by incorporating OpenAI APIs and the company's internal knowledge base.
- Getting a deeper understanding of Large Language Models
- Building framework for using OpenAI ChatGPT
- Building an efficient solution for In-Context learning based on internal knowledge and incorporating the best practices of the prompt engineering
- Implementation of chat front-end application using Svelte framework
- Applying data exploration, cleaning, and anonymization techniques (especially for PII data), preparing an efficient form of support ticket data for fine-tuning LLMs

SZILVESZTER TÓTH

+36/70 933-0910

totszilveszter@gmail.com

WORK EXPERIENCE

- Implement custom data extractor from Zendesk
- Implement ETL and machine learning pipeline in Databricks
- Deployment of the Svelte front-end and FastAPI-powered backend using GitHub actions CI technology into Azure

MOL Zrt. – PB sales management – PB gas rented filling forecasting

- Applying CRISP DM methodology on the project
- Analyzing and understanding the business domain, exploring potentially important datasets
- Elaborating deeper exploratory data analysis (EDA) on the relevant datasets with Python, Jupyter notebook and Sklearn technologies
- Creating connections between different datamarts (SAP, Excel, Worx)
- Loading data (ETL) into Hadoop Impala based data warehouse
- Creating (for the business side) useful Power BI reports as a by-product
- Applying feature engineering based on the deeper data exploration
- Modeling with ARIMA, SARIMAX, Random Forest, NN (neural network), Exponential Smoothing algorithms, and Fourier Transform algorithms
- Elaborating partial installation on the business side (data updating is supported by developers)

Zenitech Data-Driven Transformation

- Successfully addressed the challenge of rapid company growth using PowerBI, providing high-level views of the organization's current and past states, enabling management to make informed decisions for planning, forecasting, and workflow improvement.
- Collaborated with management to define and prioritize KPIs, emphasizing continuous and direct communication throughout the reporting project, with real-time dashboards powered by PowerBI.
- Developed detailed visualizations using PowerBI, including monthly, site-specific, seniority-based, and billability breakdowns, to present complex calculations and measurements for the board, uncovering utilization issues and anomalies.
- Identified and resolved company-wide issues related to projectless free capacity using PowerBI, leading to a more efficient administration flow and eliminating redundant processes.
- Developed monthly operational reports for delivery managers and people leads using PowerBI, simplifying administration processes and improving information accuracy.

SZILVESZTER TÓTH

+36/70 933-0910

totszilveszter@gmail.com

WORK EXPERIENCE

- Advocated for continuous improvement by eliminating redundant administrations and unsynchronized data sources, ensuring that PowerBI reports became a reliable and indispensable tool for management decision-making and planning.
- Emphasized the importance of communicating with stakeholders, providing diverse data representations, and deep domain understanding for effective decision support
- Defined and implemented key performance indicators (KPIs), with a focus on Billable Utilization, leading to optimized layoff strategies and improved resource allocations, all powered by the capabilities of PowerBI.
- Joining different sales pipeline stages into Power BI reports to get a way better overall picture of the company's sales pipeline and, this way ease decision making

MVM – CO2 quota trading

- Getting familiar with trading basics
- Getting familiar with applicable algorithmic trading frameworks: Freqtrade, Backtrader, Zipline
- Elaborating efficient ETL pipeline for the source data
- Implementing customized trading strategies based on CO2 quota data exploration
- Injecting ML algorithms into specific strategies
- Exploring, comparing, and using similar but much larger datasets (like cryptos)
- Handling efficiently highly frequent quote and trade datasets
- Building efficient model evaluation using MLFlow framework
- Implementing sophisticated neural networks based on publications in PyTorch framework
- Elaborating various risk-reward strategies using the model output and other tools like stop loss/take profit or stake setups
- Implementation of automated continuous retraining solution using Azure technologies and Azure ML studio
- Implementing a decoupled multi-container web socket-based solution for processing level 1 order book data, running the neural network inference, and applying the results to the trading strategy
- Implementation of a Svelte based web app for the traders

Fradi – attendance prediction

- Getting familiar with football world especially the dependence of match attendances
- Exploratory data analysis on Fradi attendance data
- Using simple feature selection mechanism with Boruta and random forest feature importances
- Building and training models (like random forest or xgboost) for predicting next matches importances based on our exploratory data analysis
- Making easy-to-integrate solution with docker composition and exposing out the service with REST interface (documented with Swagger)

SZILVESZTER TÓTH

+36/70 933-0910

totszilveszter@gmail.com

WORK EXPERIENCE

Telio – Inmate face

- Implementing efficient solution for face detection and recognition using DLIB
- Elaborating warning signals based on the result of the face recognition service
- Building easy-to-integrate architecture with Docker, RabbitMQ, Flask.
- Getting to know basics of RabbitMQ and AMQP protocol
- Implementing event driven endpoints using AMQP protocol

MOL Zrt. – MOL Track flight video analysis

- Determining most applicable deep learning based object detector networks
- Getting to deeper knowledge and training of YOLOv4 network
- Implementing basic alert system based on a simple image annotator tool
- Getting to know and applying deep learning based anomaly detectors
- Integrating simple tracker algorithms in the processing
- Implementing efficient algorithm for comparing multiple flight tracks (for detecting object difference)

Gránit Bank – identification and address card reading and processing on images

- Getting to deeper knowledge of Tesseract OCR and implementing its training and testing procedures
- Implementing efficient card detection by applying SIFT detector and its key feature points
- Participating in the elaboration of the image processing pipeline and testing environment
- Elaboration of pre- and postprocessing algorithms in order to significantly increase the accuracy and efficiency
- Implementing effective DevOps pipeline for the fast development and delivery by using Docker and Gitlab CI

Grid Data Management System – household electricity consumption forecasting

- Elaborating deeper exploratory data analysis (EDA) on the relevant datasets with Python, Jupyter notebook and Sklearn technologies
- Deeper analysis and understanding of the business and industrial (domain) area
- Collecting, analyzing, and applying household datasets from several foreign countries
- Implementing efficient algorithm for creating the most similar dataset compared to the Hungarian conditions
- Elaborating simple modeling framework
- Modeling with ARIMA, SARIMAX, Random Forest, NN (neural network), Exponential Smoothing algorithms
- Applying feature engineering on the datasets

SZILVESZTER TÓTH

+36/70 933-0910

totszilveszter@gmail.com

WORK EXPERIENCE

Hungarian post office – electronic delivery system - Reports

- Participating in implementation of data warehouse for reporting
- Creating data marts for specific reports and report groups
- Implementing and optimizing data efficient extraction, transformation, and load (ETL) from source systems
- Optimizing SQL queries, inspecting execution plans, applying appropriate indexes
- Designing star schema

Hungarian post office – electronic delivery system - Backend

- Participating in .NET backend development
- Optimizing electronic delivery system report module's performance

Machine Learning Engineer
Idaso Ltd., Ireland, www.idaso.ie 2016 FEB - 2018 DEC

Vehicle detection based on images

- Getting to know Deep Learning basics, with particular regard to convolutional neural networks
- Getting to know, train and examine neural network-based object detectors and classifiers on public and company-provided data
- Examining SSD (Single Shot Multibox Detector), YOLO (You Only Look Once), and Faster R-CNN object detectors, targeting and using for Vehicle Detection
- Getting to know, examine and use AlexNet, Inception and ResNet classification networks to categorize vehicles
- Study and use SORT (Simple Online and Realtime Tracking) and Deep SORT object trackers to track vehicles
- Creating end-to-end automated processing of vehicle counting by implementing a Serverless architecture in a Microsoft Azure environment that includes the use of previously mentioned components
- Getting to know Azure Batch Service and developing automatic processing
- Getting to know the NVIDIA Jetson TX 2 Module and running the SSD Detector and Deep SORT tracker components on the device and optimizing them for reaching the right speed (Goal: Online vehicle counter camera)
- Collecting and generating training data for object detection and classification using and modifying GTA (Grand Theft Auto) V game software and RenderDoc graphics debugger