Binh Pham

</> [Portfolio](#) • in [LinkedIn](#) • ✉ binh.pham5918@gmail.com • 📞 [(+358)414805030](#)

Experience

KONE *Helsinki, Finland*
Data Engineer Intern *May 2024 - December 2024 (8 months)*

- **Technologies:** TypeScript, Python, SQL, Terraform, AWS, Databricks Unity Catalog, dbt, Docker, Spark.
- Built reusable cross-environment Terraform modules, efficiently provisioning cloud resources including AWS Glue tables, Unity Catalog schemas, Databricks tables, and access grants.
- Maintained and debugged existing AWS Glue ETL scripts in PySpark, ensuring cost-effective performance and reliable data quality in production environment.
- Provisioned new AWS IAM resources and improved current S3 configurations using AWS CDK, leveraging fine-grained access control and optimizing storage costs.
- Developed and containerized dbt models into Docker images, orchestrated their scheduled executions with AWS Batch jobs, producing Delta Lake tables used for business intelligence.
- Designed and implemented GitLab CI/CD pipelines, automating the releases of IaC updates within data workflows.

Slush *Helsinki, Finland*
Junior Software Engineer *September 2023 - December 2023 (4 months)*

- **Technologies:** TypeScript, PostgreSQL, Express, React, React Native, Node, Google Cloud Platform, Docker.
- Developed an in-house lunch matching (LM) app to replace third-party service, saving $5,000 annually.
- Reduced the LM algorithm's time complexity from $O(n^3)$ to $O(n \cdot log(n))$, space complexity from $O(n^2)$ to $O(n)$.
- Built full-stack features with PERN stack for the administrative system used to monitor all organisation events.
- Implemented key front-end components in the Slush website and mobile app, serving 40,000 users.
- Enhanced the meetings booking feature on the mobile app, decreasing front-end latency by 3 seconds.

ABB *Helsinki, Finland*
Data Analyst Intern *January 2023 - August 2023 (8 months)*

- **Technologies:** Python, SQL, Azure, Databricks, Power BI, scikit-learn, Spark.
- Created PySpark analytics functions to extract quality metrics from IoT data and visualized insights through Power BI reports, enabling daily monitoring of production data health.
- Developed IO libraries within Databricks environments, streamlining the processes of reading from and writing to Azure Data Lake for the data science team.
- Implemented an ensemble regression model with scikit-learn, predicting data storage costs with high accuracy.

Elisa *Helsinki, Finland*
Software Engineer Intern *May 2022 - September 2022 (4 months)*

- **Technologies:** Python, SQL, Azure, Databricks, sciki-learn, Spark, TensorFlow.
- Migrated users data from a legacy information model to a refined schema, significantly enhancing accessibility.
- Built out-of-the-box data processing functionalities with Spark for various machine learning workflows.
- Developed automated machine learning solutions to accelerate the models and parameters selection process.

Education

Aalto University *Espoo, Finland*
Master of Science in Computer Science *September 2022 - May 2025 (expected)*

- Minor in Security and Cloud Computing.

Aalto University *Espoo, Finland*
Bachelor of Science in Data Science *September 2019 - June 2022*

- Minor in Computational Engineering.

Technical Skills

Languages: TypeScript (JavaScript), Python, C#, Scala, SQL, YAML, HTML/CSS

Software Engineering: PostgreSQL, MongoDB, Node, Express, Flask, React, React Native, Git, Linux, Docker

Data Engineering: Amazon Web Services (AWS), Azure, Databricks, dbt (data build tool)

Machine Learning: TensorFlow, Keras, scikit-learn, PyTorch, OpenCV