

MOHANAD ELSAFTY

Singapore, Singapore
+6588498090

mohanad.elsafty@outlook.com
[linkedin.com/in/mohanad-elsafty](https://www.linkedin.com/in/mohanad-elsafty)

EMPLOYMENT

Data Infrastructure Engineer

Shopee

Jan 2021 – Present

Apache Spark (Scala, Java, JavaScript)

- **Dramatically enhanced** Spark history server for handling millions of records effortlessly, resulting in **80% performance improvement** with larger datasets and enable search and filter on client side, **reducing** server load by **45%** and **enhancing** user experiences.

Apache Hadoop/HDFS (Java, Airflow, Kubernetes)

- **Developed** logs aggregation for Yarn/HDFS logs, **leading to more efficient monitoring and saving 3 hours weekly**.
- **Designed and implemented** a low-memory fingerprint for extended offline analysis of enormous size FSIMAGE, **allowing biggest FSIMAGE to fit within just 30% of its previous memory footprint**.
- **Engineered** a block-based InputFormat for DistCp (MapReduce application), **optimizing** workload distribution among mappers **based on blocks count exploiting data locality and shortening data transfer time by 20%**.
- **Enhanced** DFSClient and NameNode to effectively mitigate the impact of sluggish DataNodes on read/write operations.

Apache Ozone (Java, C++, Golang, JavaScript)

- **Engineered** a sophisticated petabyte scale data Retention Service akin to AWS S3 lifecycle configurations saving 15% of wasted storage and created AWS CLI/SDK compatible s3gateway endpoint for easier integration with existing clients.
- **Architected** a universal s3gateway solution, **empowering user access to multiple clusters** without compromising performance or incurring additional latency.
- **Devised** an efficient fair call queue for s3gateway, **ensuring fairness for all clients**. Additionally, Applied **Quality of Service mechanisms** for s3gateway, **safeguarding impeccable service delivery while preventing any misuse or exploitation by clients**.

Data Engineer

Global Fashion Group

Jan 2019 – Dec 2020

ZMM (Markdown recommendation system) (Airflow, Python, SQL, Pandas/NumPy, Notebooks)

- **Managed the process of revamping and refactoring** the existing system into python and Jupyter notebook, **employed containerization, restructured, and fine-tuned** ZMM. These efforts enabled **efficient and auto-scalable deployment and cost reduction**, resulting in an **impressive 41% reduction in runtime** and **70.6% cost savings**.
- **Optimized SQL queries** using the latest features of Amazon Redshift **slashing ingestion time by 66%**. **Streamlined scheduling and deployment end-to-end** through **Apache Airflow**, [ECS](#), [CodeBuild](#).

ZMA (Media attribution model) (Python, R, SQL, Pandas/NumPy)

- **Pioneered a highly effective media attribution model based on Hidden Markov model**, improving marketing decision-making by **20%**.

[Github](#)

[Bitbucket](#)

[LeetCode](#)

[Hackerrank](#)

[Linkedin](#)

- Collaborated with management to visualize model outcomes using Apache Superset for informed decision-making.

CPA (Automated prices prediction) (PySpark, Python, pandas, Catboost model)

- Led the development, enhancement, and optimization of an automated price prediction system.
- Designed a novel data pipeline, leveraging Pyspark SQL interface and built-in ML features improving training time by 25% over existing pipeline.
- Devised and implemented highly efficient data ingestion techniques, data bridges and optimizing cluster size and operational costs.

Full Stack Developer

SigmaProIT

Dec 2017 – Sep 2018

Experto (App to App VOIP connects clients with experts) (Python, Django, NodeJS, Angular/Ionic)

- Developed a VOIP android application with multiple authentication methods along with statistical views and dedicated webservice to keep track of voice calls lifecycle.

Ziruna (Expert booking app makes it easy for consumer to hire providers) (Python, Django, Angular/Ionic)

- Participated in mobile application development and optimized restful backend API response time by 20%.

Software Engineer

Clouddypedia

Apr 2016 – Nov 2017

(Python, Django, GoogleAppEngine, ElasticSearch, OpenEDX)

- Constructed a certification module to support Arabic language for enhanced accessibility and customized the registration module to integrate specific OAuth providers and domains.
- Built filter and tagging module to optimize search capabilities on OpenEDX search module reducing search time by 30%.
- Designed a googleAppEngine based web application slashing email migration time by 40% and crafted python wrappers for G-Suite APIS to enhance application integration.

EDUCATION

Cairo, Egypt

Ain Shams University

2011 – 2015

- B.Sc. in Computer Science and information systems (Scientific computing department).

Languages and Technologies

- Python; Java; Scala; Golang; JavaScript; C++; Apache Spark; Apache Hadoop; Apache Ozone; RocksDB; Apache Kafka; SQL; PostgreSQL; Django; Flask; Webapp2; Angular; ReactJS; Ionic; AWS services; GoogleAppEngine; Heroku; Docker; Firebase; Apache Hadoop; Apache Ozone; Apache Spark; Distributed systems; Pandas; NumPy

[Github](#)

[Bitbucket](#)

[LeetCode](#)

[Hackerrank](#)

[Linkedin](#)