

Σοφία Αργυρίου

PySpark Data Engineer

✉ σοφία.αργυρίου@example.com

☎ +30 69 412 71 20

📍 Τρίκαλα



Περίληψη

PySpark Data Engineer με εστίαση σε επεξεργασία αισθητήρων ρομπότ. Δημιουργώ αξιόπιστα data products με PySpark, Delta Lake, Structured Streaming, Kafka, MLflow και συνδέω τις τεχνικές επιλογές με μετρήσιμα αποτελέσματα λειτουργίας.



Εργασιακή Εμπειρία

PySpark Data Engineer

Jan 2022 - Παρόν

Θεσσαλική Ρομποτική · Τρίκαλα

- Σχεδίασα pipelines με PySpark και Delta Lake για επεξεργασία αισθητήρων ρομπότ στην Θεσσαλική Ρομποτική, σε συνεργασία με product και platform teams για 3.200 ρομπότ; αύξησα την έγκαιρη παράδοση από 92% σε 99,6%.
- Ανακατασκεύασα μοντέλα Structured Streaming και partitioning για το έργο «Feature lake για ρομποτικούς αισθητήρες», συντονίζοντας ομάδα 9 μηχανικών· μείωσα τον median χρόνο ερωτήματος από 24 σε 4,1 δευτερόλεπτα.
- Εγκατέστησα ελέγχους schema, freshness και lineage σε 38 production tables για επεξεργασία αισθητήρων ρομπότ, με 14 αναλυτές ή application teams· περιόρισα τα data incidents κατά 63%.
- Αυτοματοποίησα backfills, alerts και runbooks γύρω από Kafka jobs στην Τρίκαλα, υποστηρίζοντας 12 αγορές ή επιχειρησιακές μονάδες· εξοικονόμησα 86 ώρες on-call ανά τρίμηνο.

Data Platform Engineer

Apr 2019 - Dec 2021

Θεσσαλική Ρομποτική Data Lab · Τρίκαλα

- Σχεδίασα pipelines με PySpark και Delta Lake για επεξεργασία αισθητήρων ρομπότ στην Θεσσαλική Ρομποτική, σε συνεργασία με product και platform teams για 3.200 ρομπότ; αύξησα την έγκαιρη παράδοση από 92% σε 99,6%.
- Ανακατασκεύασα μοντέλα Structured Streaming και partitioning για το έργο «Feature lake για ρομποτικούς αισθητήρες», συντονίζοντας ομάδα 6 μηχανικών· μείωσα τον median χρόνο ερωτήματος από 24 σε 4,1 δευτερόλεπτα.
- Εγκατέστησα ελέγχους schema, freshness και lineage σε 38 production tables για επεξεργασία αισθητήρων ρομπότ, με 5 αναλυτές ή application teams· περιόρισα τα data incidents κατά 63%.
- Αυτοματοποίησα backfills, alerts και runbooks γύρω από Kafka jobs στην Τρίκαλα, υποστηρίζοντας 4 αγορές ή επιχειρησιακές μονάδες· εξοικονόμησα 86 ώρες on-call ανά τρίμηνο.



Έργα

Feature lake για ρομποτικούς αισθητήρες

Ανάπτυξα τεκμηριωμένο data product για επεξεργασία αισθητήρων ρομπότ, με ελεγμένους μετασχηματισμούς και μοντέλο έτοιμο για ανάλυση.

PySpark

Delta Lake

Structured Streaming

Kafka



Δεξιότητες

Τεχνικό stack

PySpark

Delta Lake

Structured Streaming

Kafka

MLflow

Παράδοση

Μοντελοποίηση δεδομένων

Ποιότητα δεδομένων

CI/CD

Observability

Τεκμηρίωση

Συνεργασία

Workshops απαιτήσεων

Incident response

Mentoring

Agile παράδοση



Εκπαίδευση

Ελληνικό Μεσογειακό Πανεπιστήμιο

Sep 2015 - Jun 2019

Πτυχίο Πληροφορικής at Data engineering και λογισμικό



Πιστοποιήσεις

- Databricks Associate Developer for Spark** May 2024
Διεθνές συμβούλιο πιστοποίησης
- Data Reliability Workshop** Aug 2023
Data Engineering Guild



Γλώσσες

Ελληνικά • Μητρική

Αγγλικά • Επαγγελματική χρήση