Combining Network Data Analytics Function and Machine Learning for Abnormal Traffic Detection in Beyond 5G

Abdelkader MEKRACHE, Karim BOUTIBA, Adlen KSENTINI

EURECOM, Sophia Antipolis, France
OpenAirInterface, Sophia Antiplis, France

Agenda

Introduction

→ 5G Networks

- 5G is powered by use cases like ultra-fast broadband, mission-critical IoT, and low-latency services.
- Key components in 5G architecture include a Radio Access Network (RAN), and a Core Network (CN)
- Standardization for 5G is carried out by several international organizations, including: ITU, IEEE, 3GPP

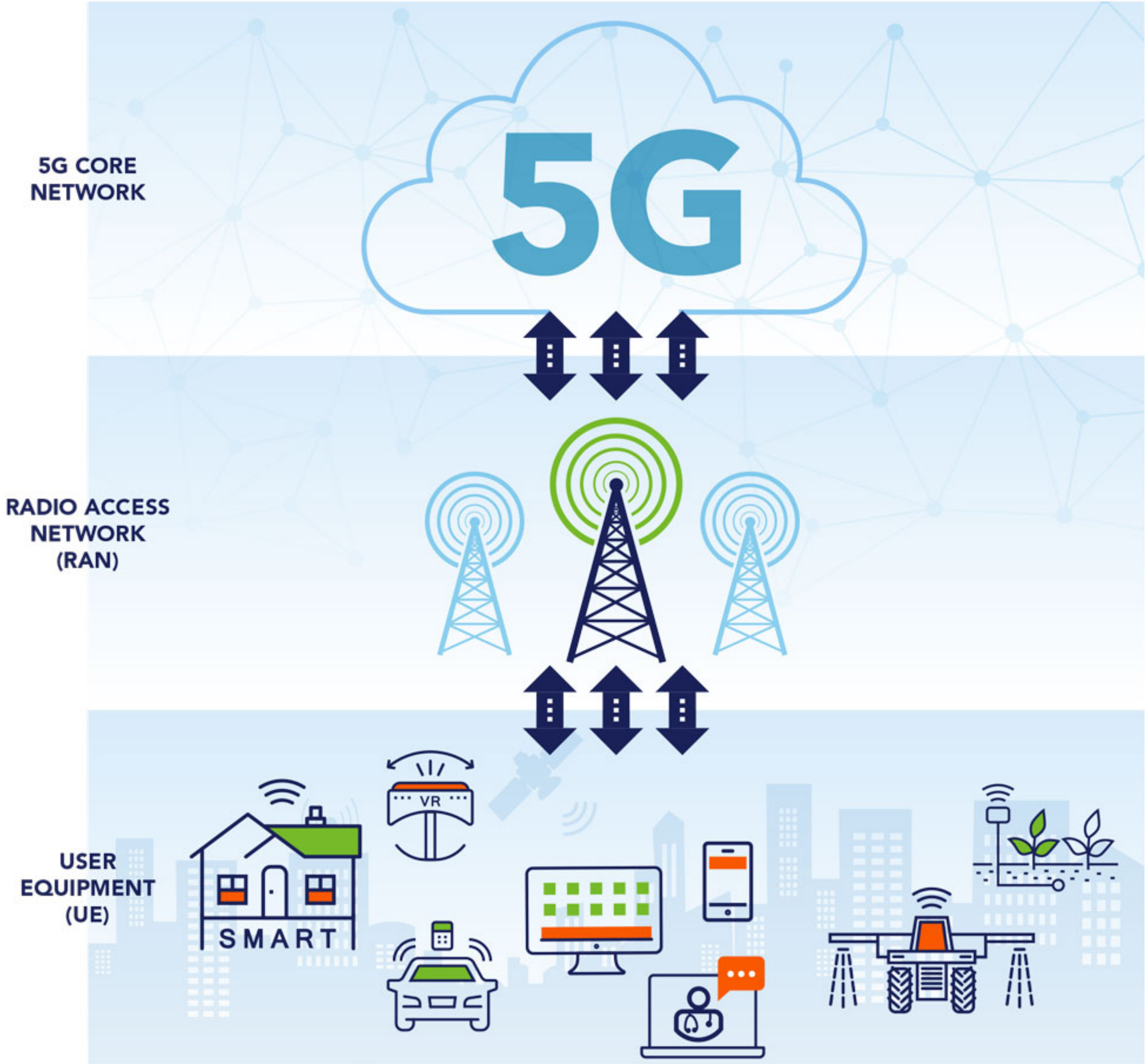


Fig. 1. End-to-End Overview of 5G Networks Architecture [1].

[1] https://software.org/wp-content/uploads/softwareorg_5Gsoftware.pdf

Introduction

→ 5G Service Based Architecture (SBA)

- 3GPP Introduced the 5G Core Network Architecture as SBA [TS 23.501].
- Communication is realized through REST API calls using the HTTP routing mechanism over TCP.
- Several NFs are defined : SMF, AMF, UDR, …, NWDAF.

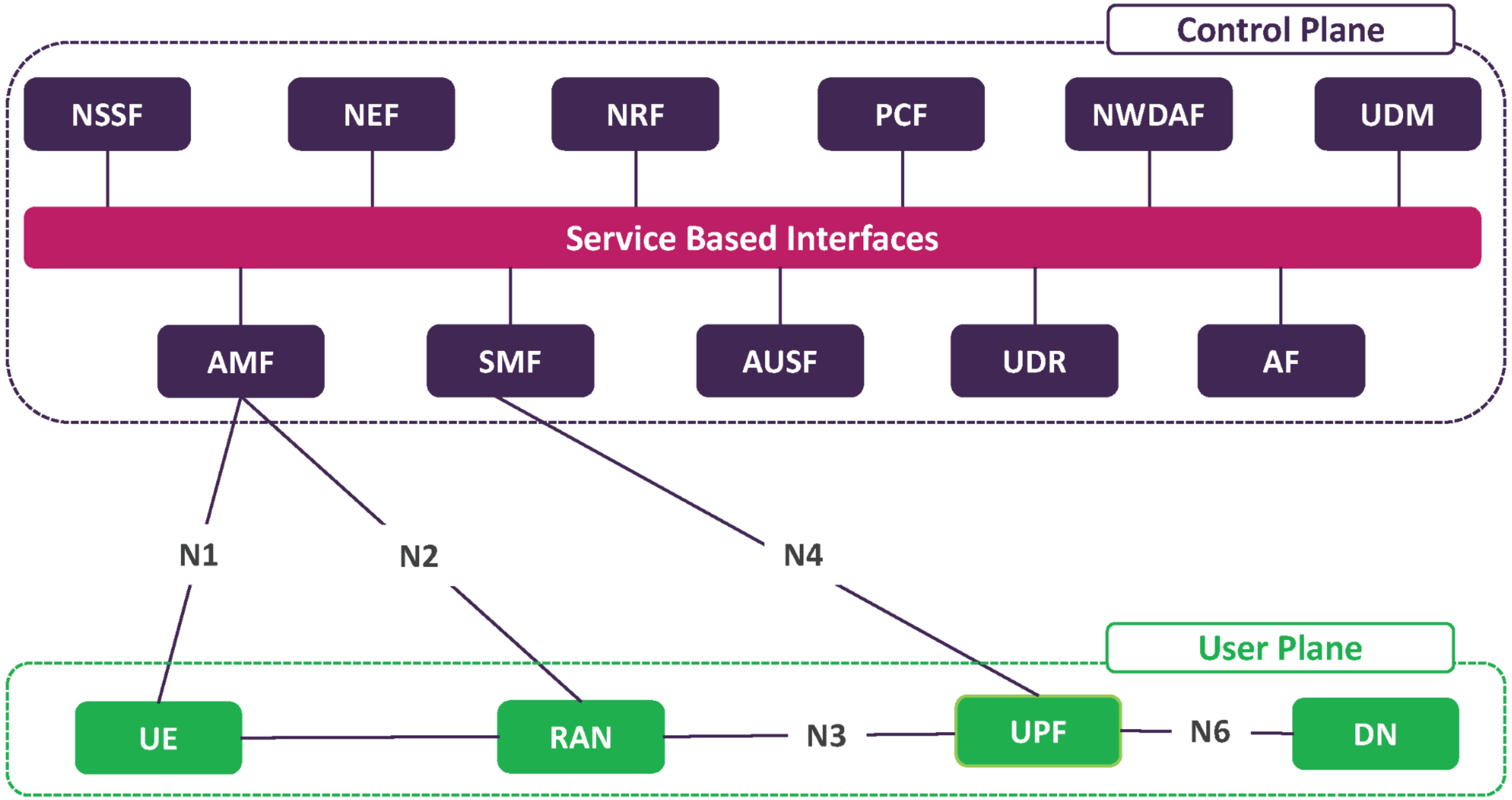


Fig. 2. 5G Service Based Architecture [2].

[2] Malik, S., Khan, M. A., El-Sayed, H., Khan, J., & Ullah, O. (2022). Implanting Intelligence in 5G Mobile Networks—A Practical Approach. Electronics, 11(23), 3933.

What is NWDAF ?

NWDAF was defined to provide analytics to 5G NFs and OAM.

NWDAF provides two main services [3]:

- The NBI Analytics Info module allows clients to request and receive specific types of analytics information
- The NBI Events Subscription module enables clients to subscribe to or unsubscribe from various analytics events.

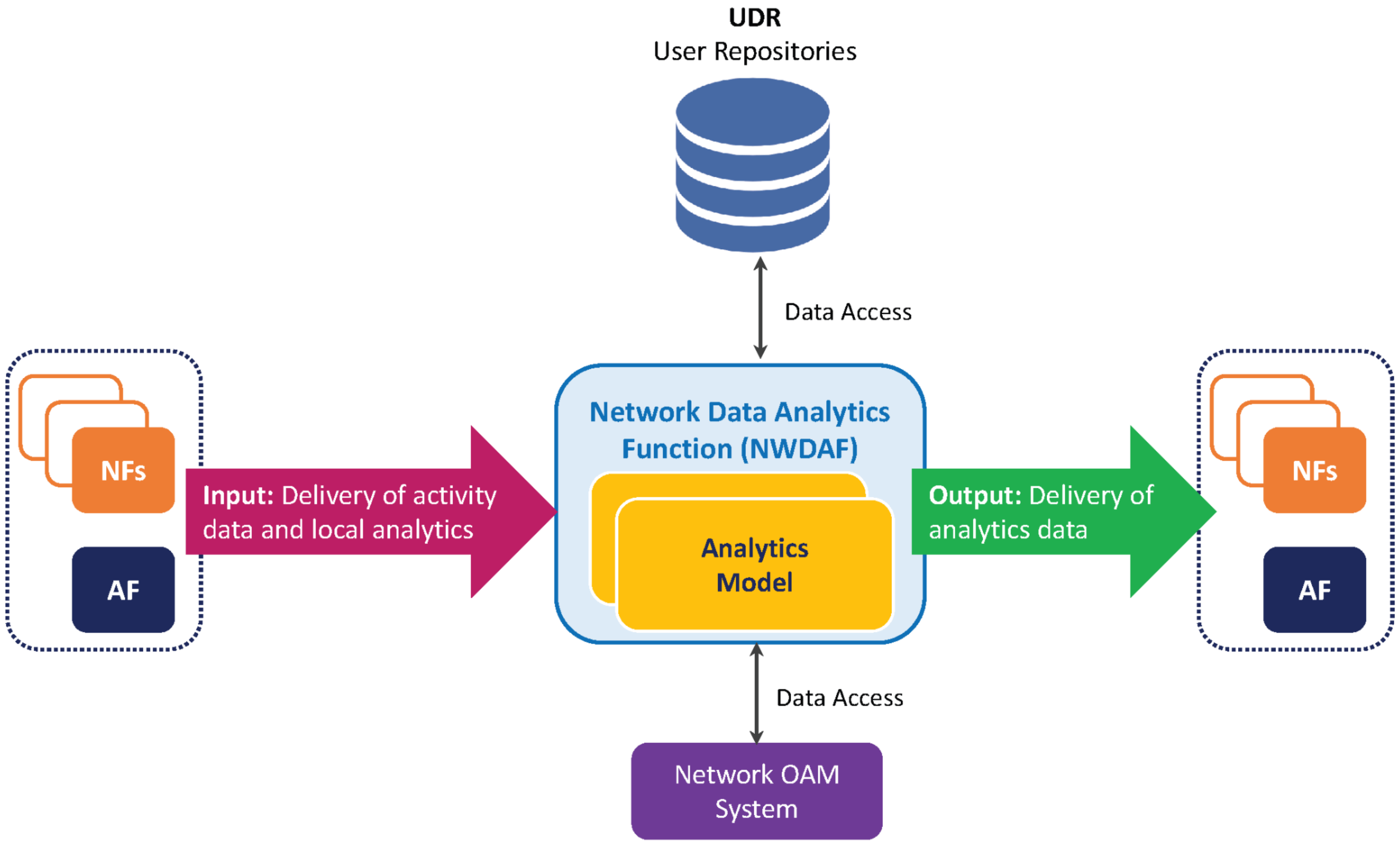


Fig. 3. Network Data Analytics Function [2].

[2] Malik, S., Khan, M. A., El-Sayed, H., Khan, J., & Ullah, O. (2022). Implanting Intelligence in 5G Mobile Networks—A Practical Approach. Electronics, 11(23), 3933.

[3] 3GPP. Technical Specification Group Core Network and Terminals; 5G System; Network Data Analytics Services; tech. rep. 29.520. Version 17.10.0. 2023

NWDAF design

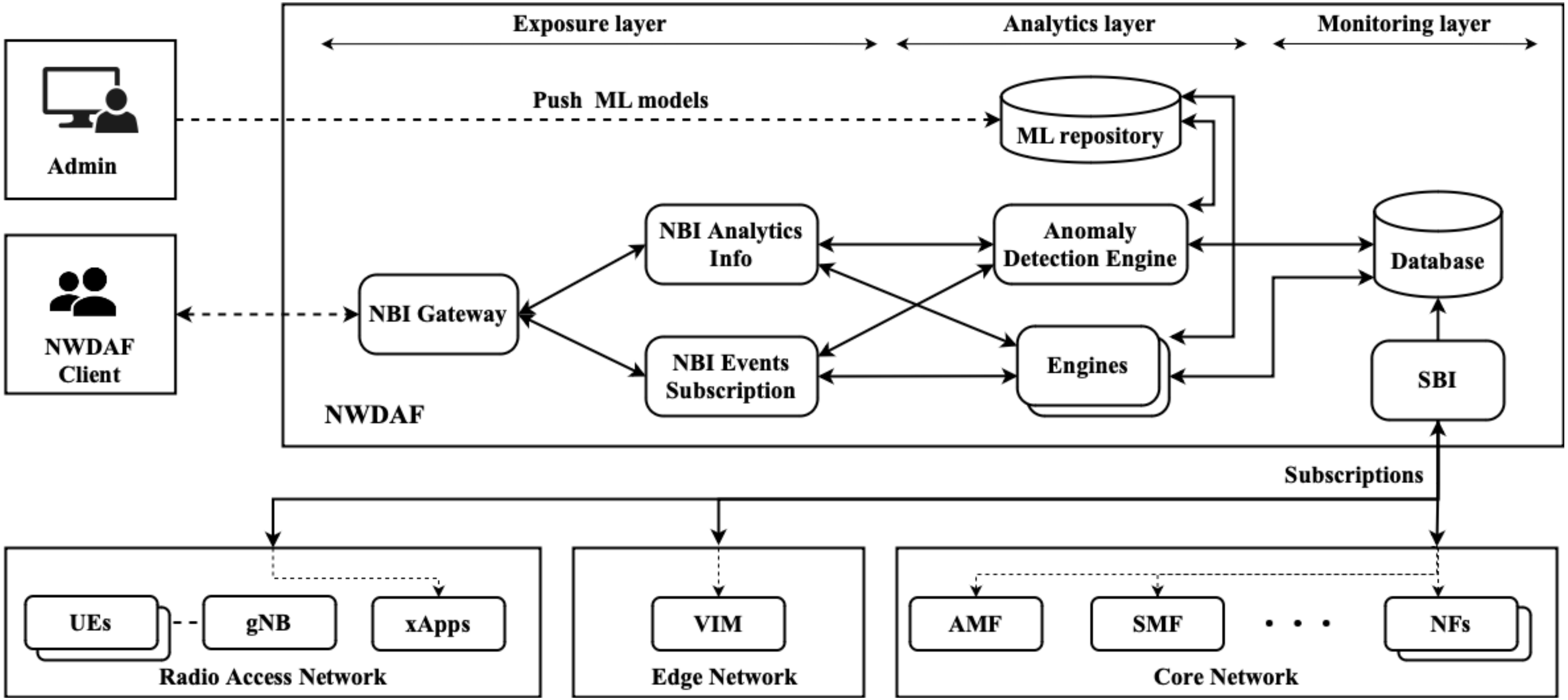


Fig. 4. Microservices-based NWDAF architecture.

NWDAF implementation

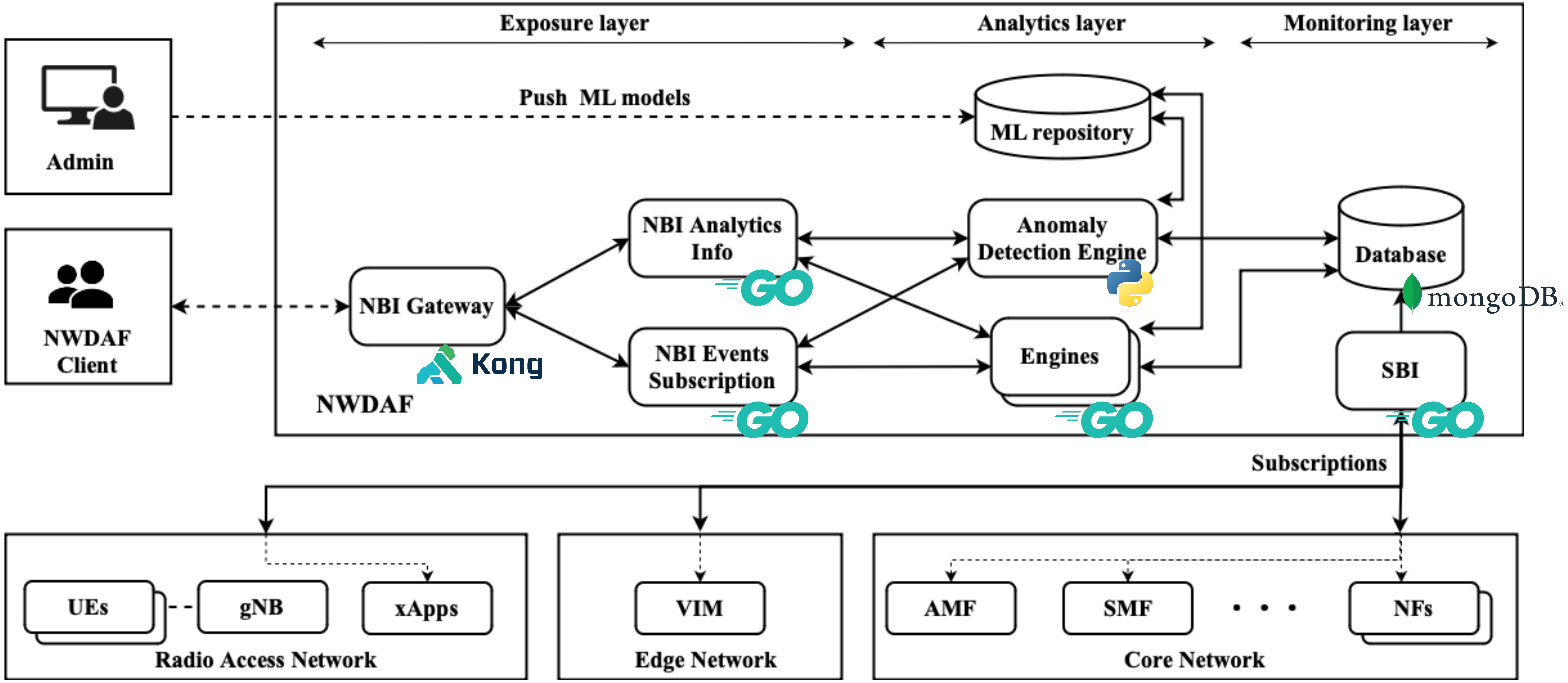


Fig. 4. Microservices-based NWDAF architecture.

NWDAF use cases

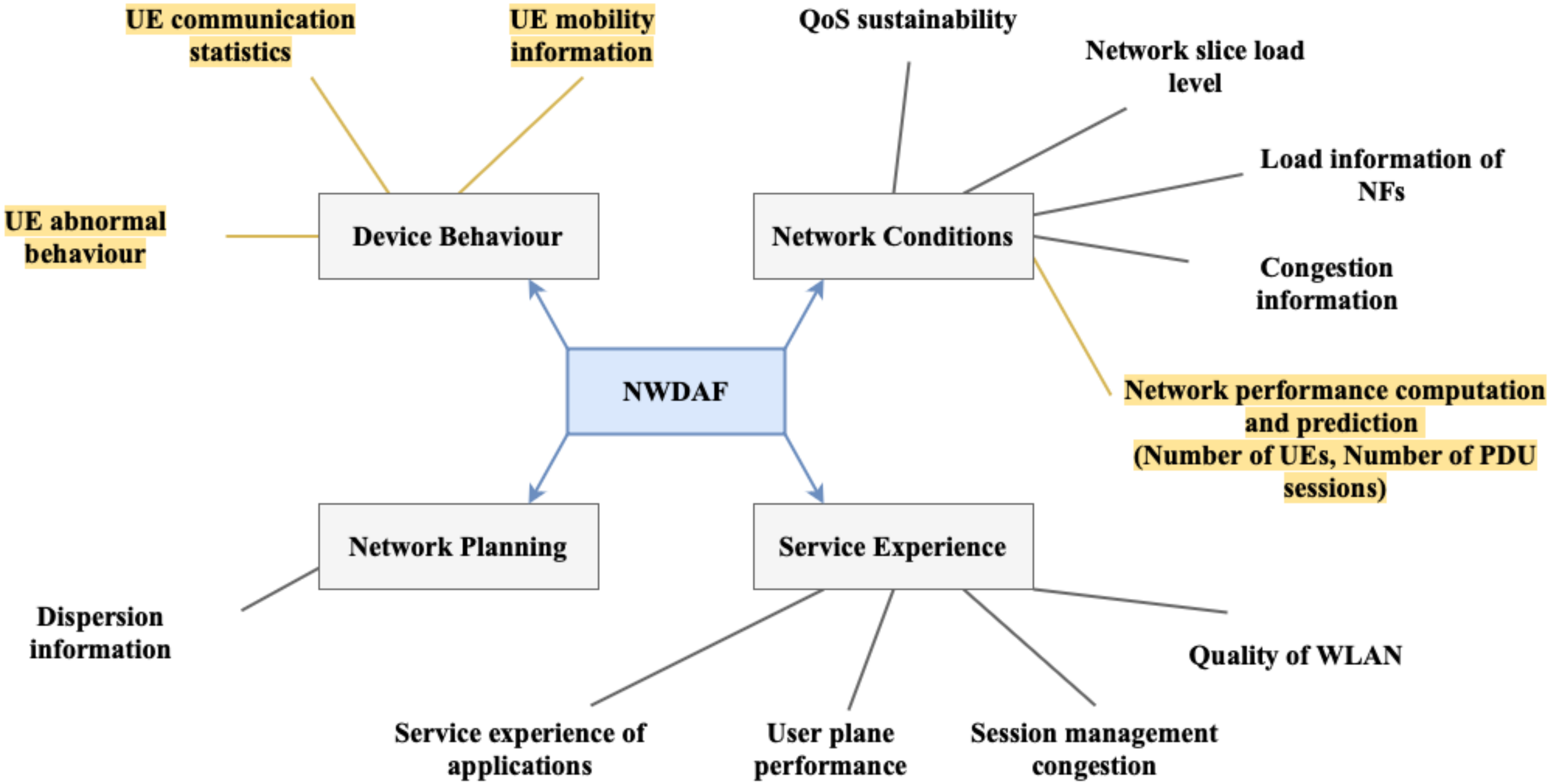


Fig. 5. A classification of 3GPP defined NWDAF Events[3].

[3] 3GPP. Technical Specification Group Core Network and Terminals; 5G System; Network Data Analytics Services; tech. rep. 29.520. Version 17.10.0. 2023

NWDAF use cases

→ Core services

- The NWDAF provides support for two types of network performance events:
  - “num_of_ue”, measures the number of attach requests during a time window.
  - “sess_succ_rate”, measures the session success rate during a time window specified in the request.
- UE Communications: “ue_comm” refer to the number of packets and bytes exchanged in the uplink and downlink directions for each PDU session.
- UE Mobility: "ue_mobility" refers to the cell ID of a given UE.

NWDAF use cases

→ ML-based services

- Abnormal Traffic: the NWDAF clients can subscribe to the "abnormal_behaviour" event and exception ID “unexpected_large_rate_flow” to receive periodic updates on the probability of abnormal traffic.

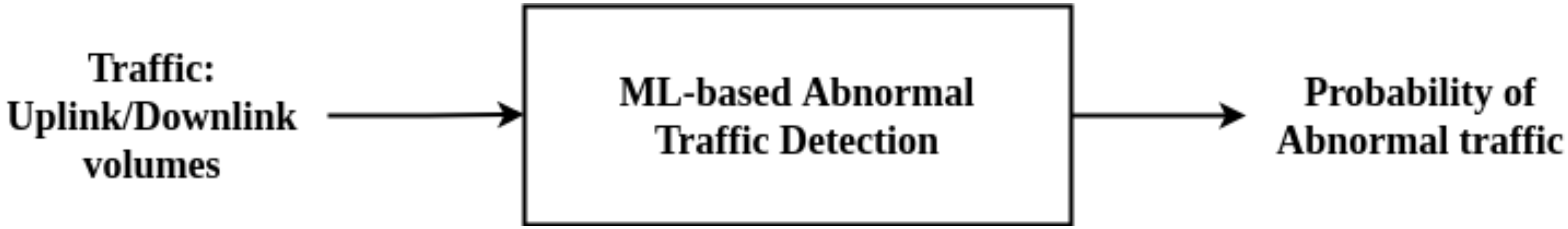


Fig. 6. ML-based Abnormal Traffic Detection model.

NWDAF use cases

→ ML-based services

- Abnormal Traffic: the NWDAF clients can subscribe to the "abnormal_behaviour" event and exception ID “unexpected_large_rate_flow” to receive periodic updates on the probability of abnormal traffic.

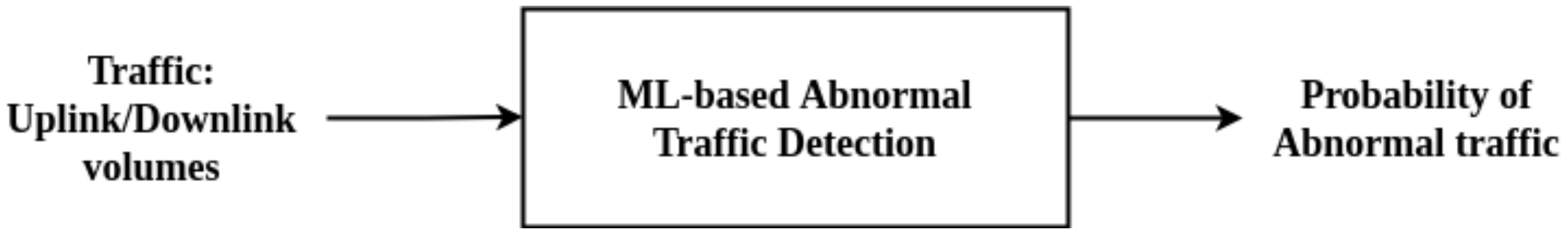


Fig. 6. ML-based Abnormal Traffic Detection model.

- Data pattern differs from UE to other → no generalization on data pattern → no labeled dataset.

- The ML-model will learn only the normal Uplink/Downlink pattern from historical data
  → It then will detect abnormal traffic → AutoEncoder Architecture.

NWDAF use cases

→ ML-based services

- Encoding: X is compressed into a lower-dimensional space $Z = \sigma(WX + b)$
- Decoding: decode Z to X' that is similar to the input dimension $X' = \sigma'(W'Z + b')$
- The loss function is $\mathrm{MAE} = \frac{\sum_{i=1}^{n} |x'_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}$
- The traffic anomaly probability is

$$p = \min(\alpha \times |\beta' - \beta|, 1)$$

α controls the impact of the distances scale.
β is the average train MAE.
β' is the test MAE.

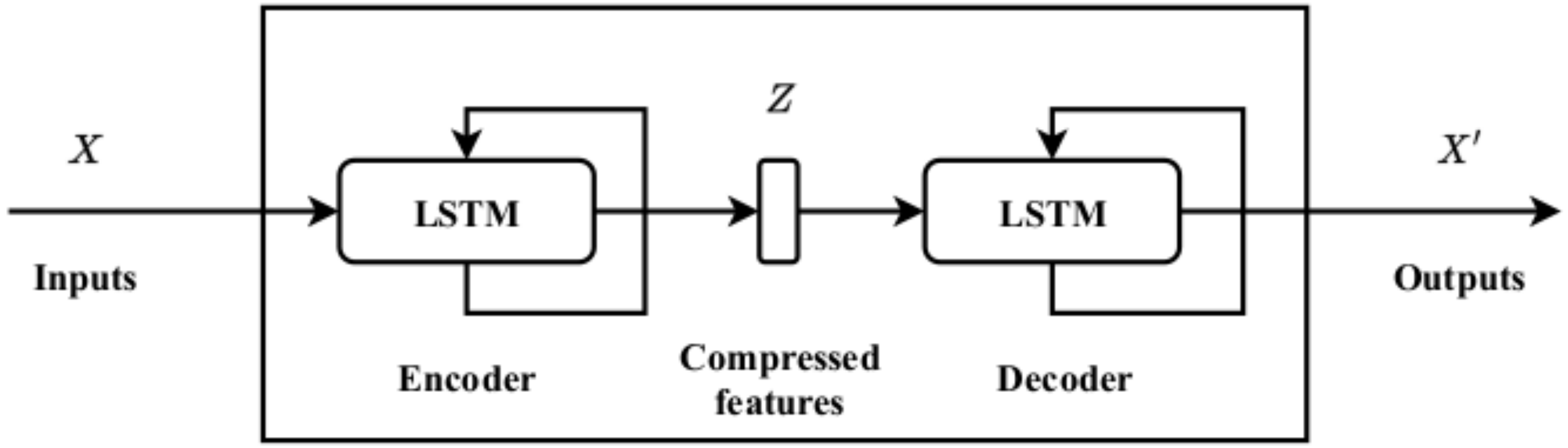


Fig. 7. LSTM Autoencoder architecture.

NWDAF use cases

→ ML-based services

- We trained LSTM-based AE using Milano Dataset [5].
- We injected some anomalies in the pattern by introducing long traffic flows.

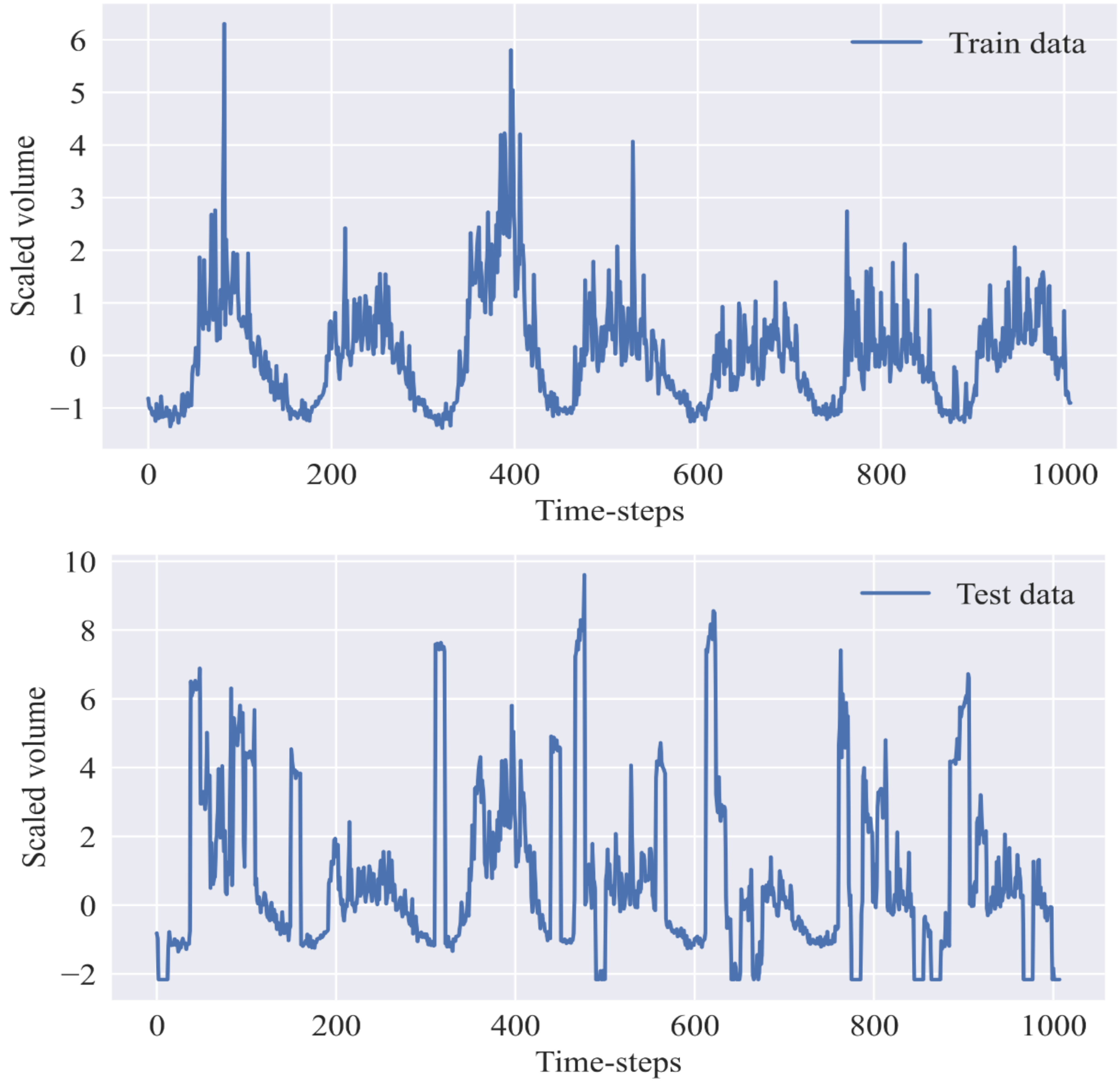


Fig. 8. Train and Test data.

[4] Barlacchi, G., De Nadai, M., Larcher, R., Casella, A., Chitic, C., Torrisi, G., ... & Lepri, B. (2015). A multi-source dataset of urban life in the city of Milan and the Province of Trentino. Scientific data, 2(1), 1-15.

NWDAF use cases

→ ML-based services

- The distance between input and generated data correlates with the anomaly probability.
- The probabilities increase as the generated data diverges from the input data.
- The anomaly probability threshold is set to 0.5

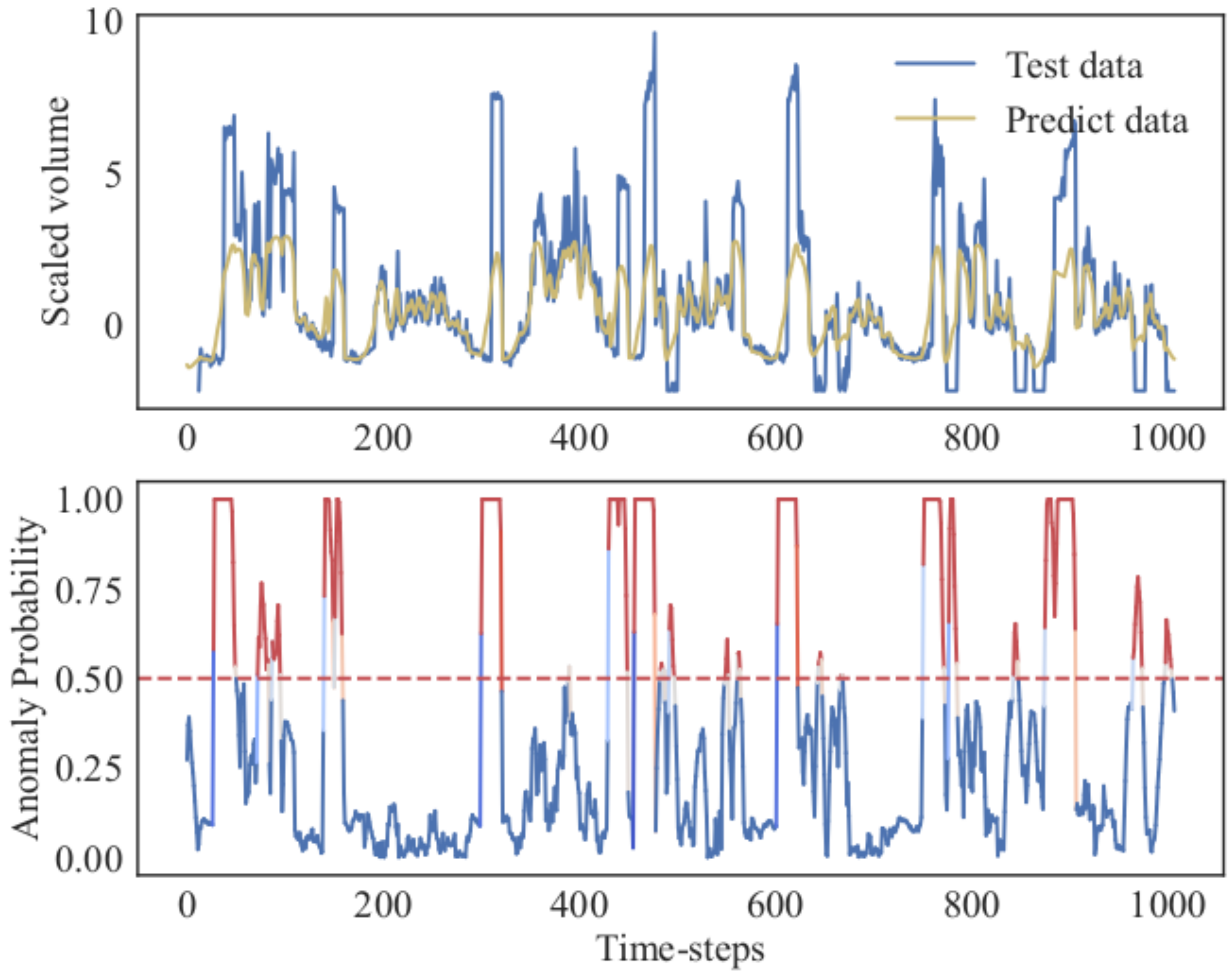
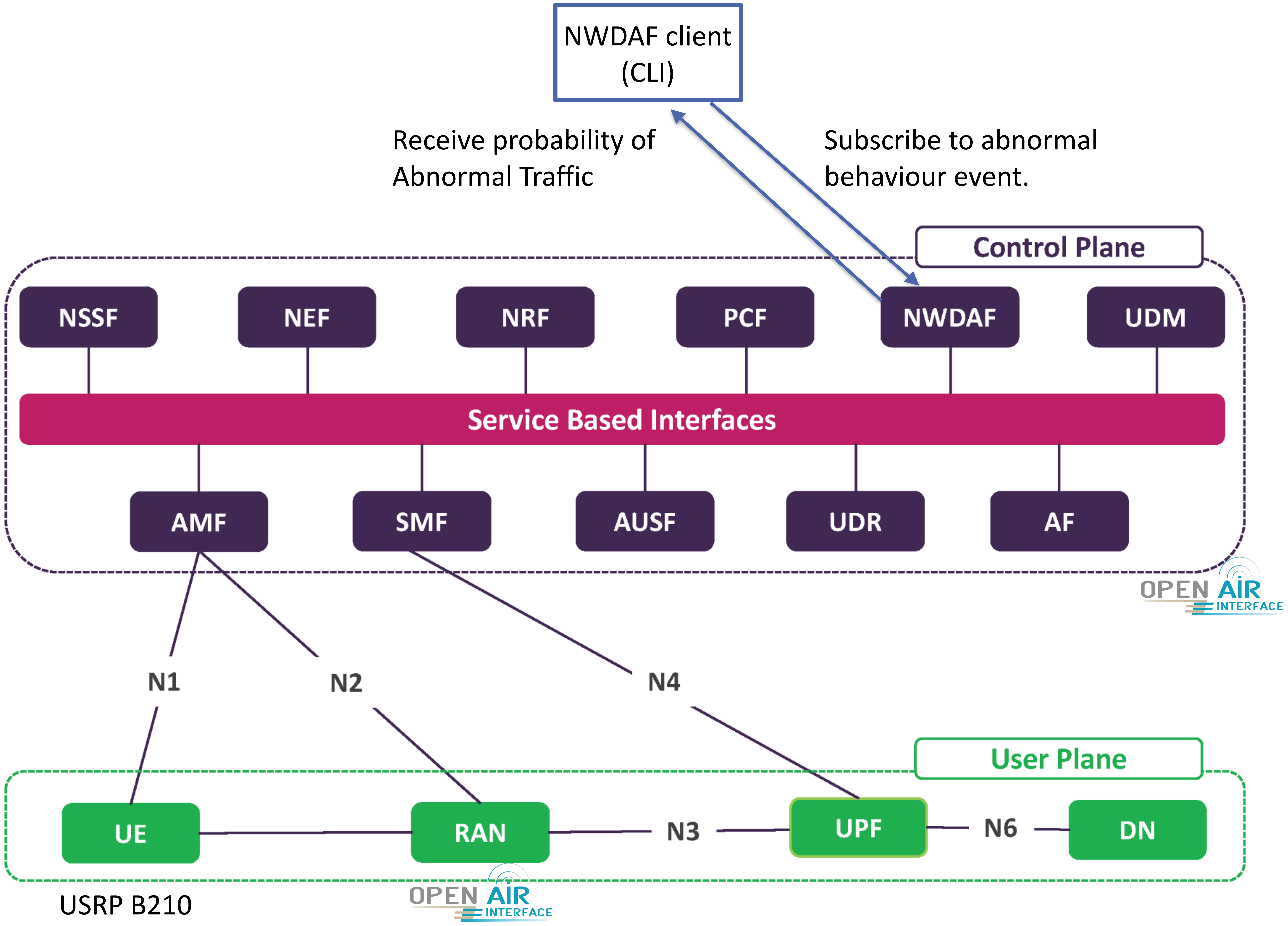


Fig. 11. Traffic anomaly probabilities for one week.

DEMO



NWDAF client (CLI)
Receive probability of Abnormal Traffic
Subscribe to abnormal behaviour event.
Control Plane
NSSF
NEF
NRF
PCF
NWDAF
UDM
Service Based Interfaces
AMF
SMF
AUSF
UDR
AF
OPEN AIR INTERFACE
N1
N2
N4
User Plane
UE
RAN
N3
UPF
N6
DN
USRP B210
OPEN AIR INTERFACE

DEMO

Thank you for your attention !
Q/A