**Advanced Optimization Algorithm Specifications**

**Title:** *Machine Learning and AI-Based Optimization Algorithm Specification for Smart 5G*

**A. Objective**

Define the architecture, data inputs, training methodology, and hyperparameters for ML models that optimize network parameters such as power allocation, beamforming, traffic steering, and predictive maintenance.

**B. Algorithm Overview**

| **Model Type** | **Purpose** | **Input Features** | **Output** | **Key Metrics** |
| --- | --- | --- | --- | --- |
| **Reinforcement Learning (RL)** | Dynamic power and resource allocation | Current network load, user QoS metrics, interference levels | Optimal power settings per cell/sector | Cumulative reward, convergence rate |
| **Supervised Learning (Random Forest, XGBoost)** | Predictive maintenance and fault detection | Historical fault logs, sensor data, telemetry time-series | Fault likelihood score | Precision, recall, F1-score |
| **Deep Neural Networks (DNN)** | Traffic pattern prediction and anomaly detection | Network KPIs, user behavior metrics, time, location | Forecasted traffic load, anomaly alerts | Mean Absolute Error (MAE), ROC-AUC |
| **Clustering (K-Means, DBSCAN)** | User segmentation for slice customization | User mobility patterns, device type, application usage | User clusters | Silhouette score, Davies-Bouldin Index |

**C. Model Architectures & Hyperparameters**

| **Model** | **Architecture Details** | **Hyperparameters** | **Training Data Size** | **Training Frequency** |
| --- | --- | --- | --- | --- |
| RL (DQN variant) | 3-layer fully connected network, ReLU activation | Learning rate: 0.001, gamma: 0.95, epsilon decay: 0.99 | 1 million episodes | Continuous, online training |
| Random Forest | 100 trees, max depth 30 | Max features: sqrt, min samples leaf: 5 | 100K labeled samples | Weekly batch training |
| DNN | 5 hidden layers, 256 neurons each, dropout 0.3 | Batch size: 128, epochs: 50, optimizer: Adam | 500K KPI samples | Daily retraining |
| K-Means | k=8 clusters, Euclidean distance | Max iterations: 300 | 1 million user sessions | Monthly re-clustering |

**D. Data Pipeline**

* **Ingestion:** Real-time telemetry via Kafka streams
* **Preprocessing:** Normalization, outlier removal, feature engineering
* **Training Environment:** GPU clusters on cloud (e.g., AWS SageMaker or Azure ML)
* **Model Validation:** Cross-validation and hold-out test sets
* **Deployment:** Containerized microservices with REST APIs

1. uning for incident response.