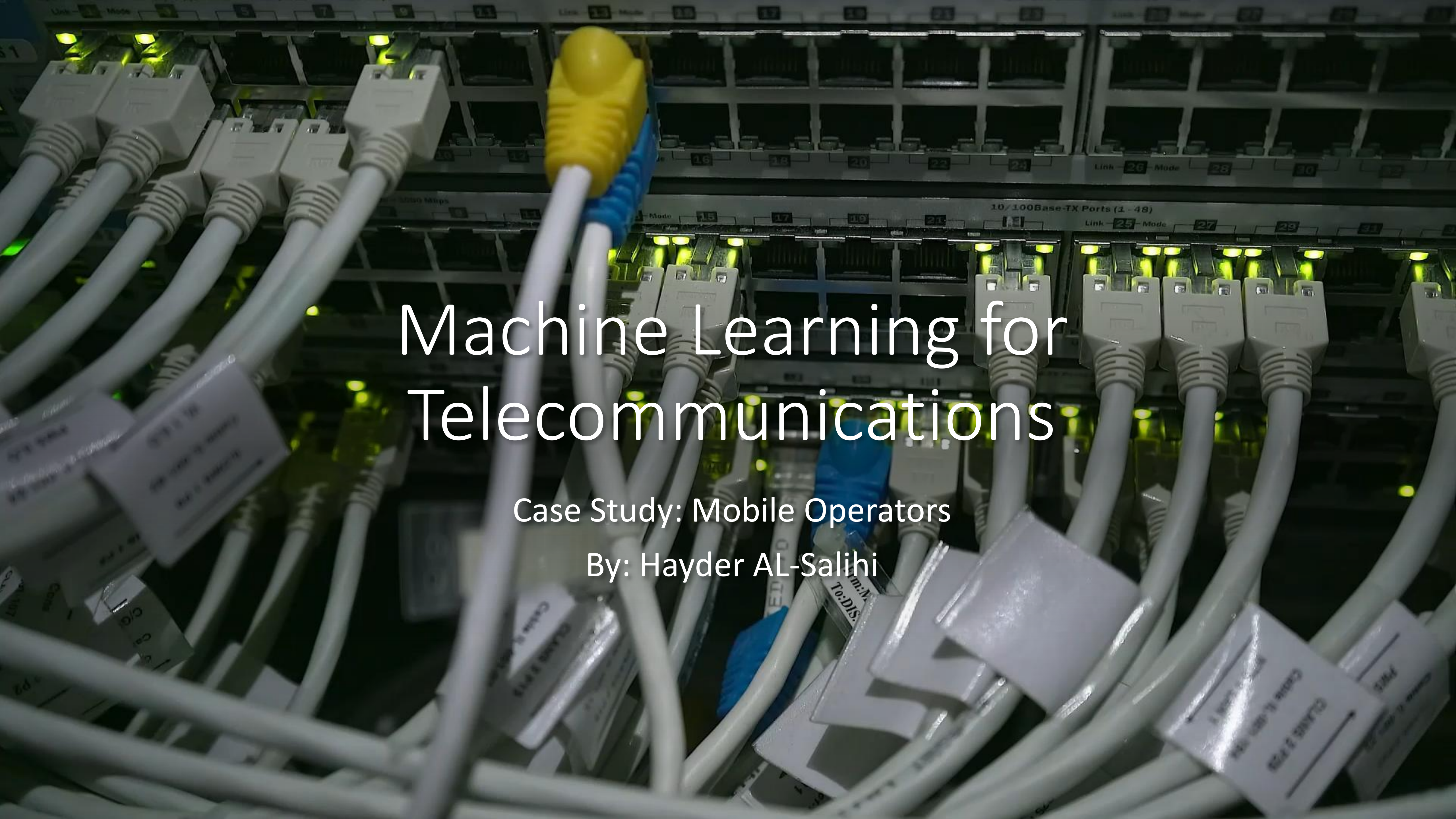


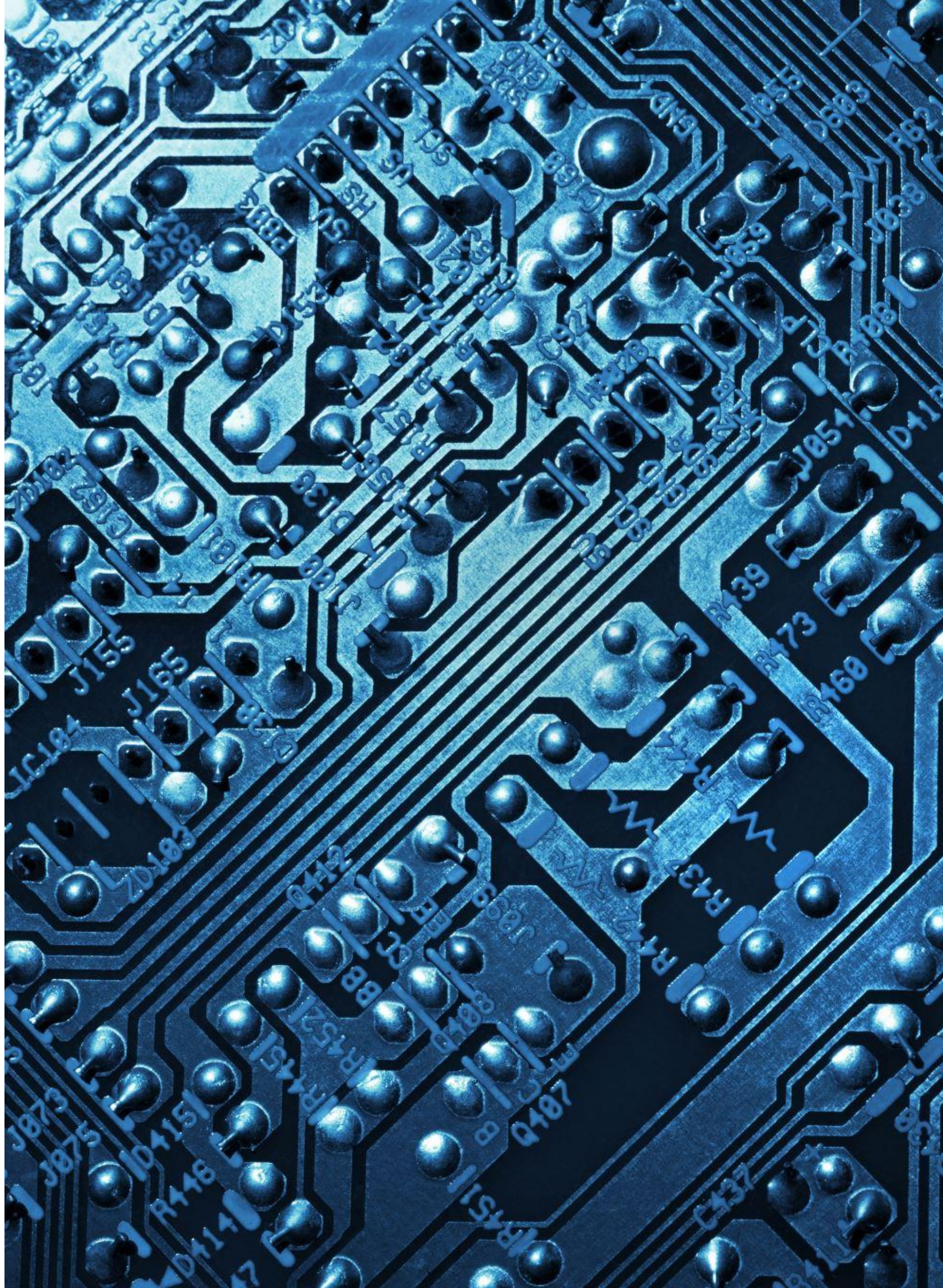
Machine Learning for Telecommunications

Case Study: Mobile Operators

By: Hayder AL-Salihi

What is Machine Learning (ML)?

- Machine Learning is a branch of **Artificial Intelligence (AI)** where:
- Computers **learn from data**
- Instead of being programmed step-by-step
- Models discover **patterns**, **relationships**, and **trends** automatically
- 👉 In telecom, this means learning from:
- Signal strength
- Dropped calls
- Customer complaints
- Usage patterns



📡 ML in a Telecom Context (Simple Flow)

Collect data (network KPIs, customer behavior, QoS/QoE)

Feed data into an ML model

Model learns patterns

Model produces outcomes (good performance / bad performance / risk)

Types of Machine Learning Used in Telecom

Unsupervised Learning (Clustering)

No labels (no “good” or “bad” given in advance)

The model groups data by **similarity**

Telecom example:

Strong signal areas → one cluster (blue)

Weak signal areas → another cluster (red)

This helps operators:

Detect **coverage holes**

Identify **problematic cells or regions**

Supervised Learning (Prediction)

Data **has labels**

Model learns from past examples

Telecom example:

Past customers → *Stayed* or *Left*

Model learns patterns linked to:

- Poor signal
- High complaints
- Low data speed

➡️ Then predicts:

“This customer is likely to churn”

The Golden Rule: Data Quality Matters

ML is only as good as the data you give it

If data is:

- Messy ❌
- Incomplete ❌
- Inaccurate ❌

Then:

Predictions will be **wrong**

Decisions will be **misleading**

✅ Clean, high-quality data = reliable ML outcomes

Key ML Use Cases in Telecom

Machine Learning enables operators to:

- 🔄 **Predict customer churn**
- 📶 **Optimize network resources**
- 📊 **Forecast traffic demand**
- 🚨 **Detect fraud**
- ⚙️ **Fix network issues faster**
- 😊 **Improve customer experience**

Predict Customer Churn

What it means:

ML analyzes customer behavior and network experience to predict **who is likely to leave**.

Uses data like:

- Call drops
- Slow internet
- Complaints
- Billing issues
- Usage decline

Why it matters:

Retaining customers is **cheaper than acquiring new ones**

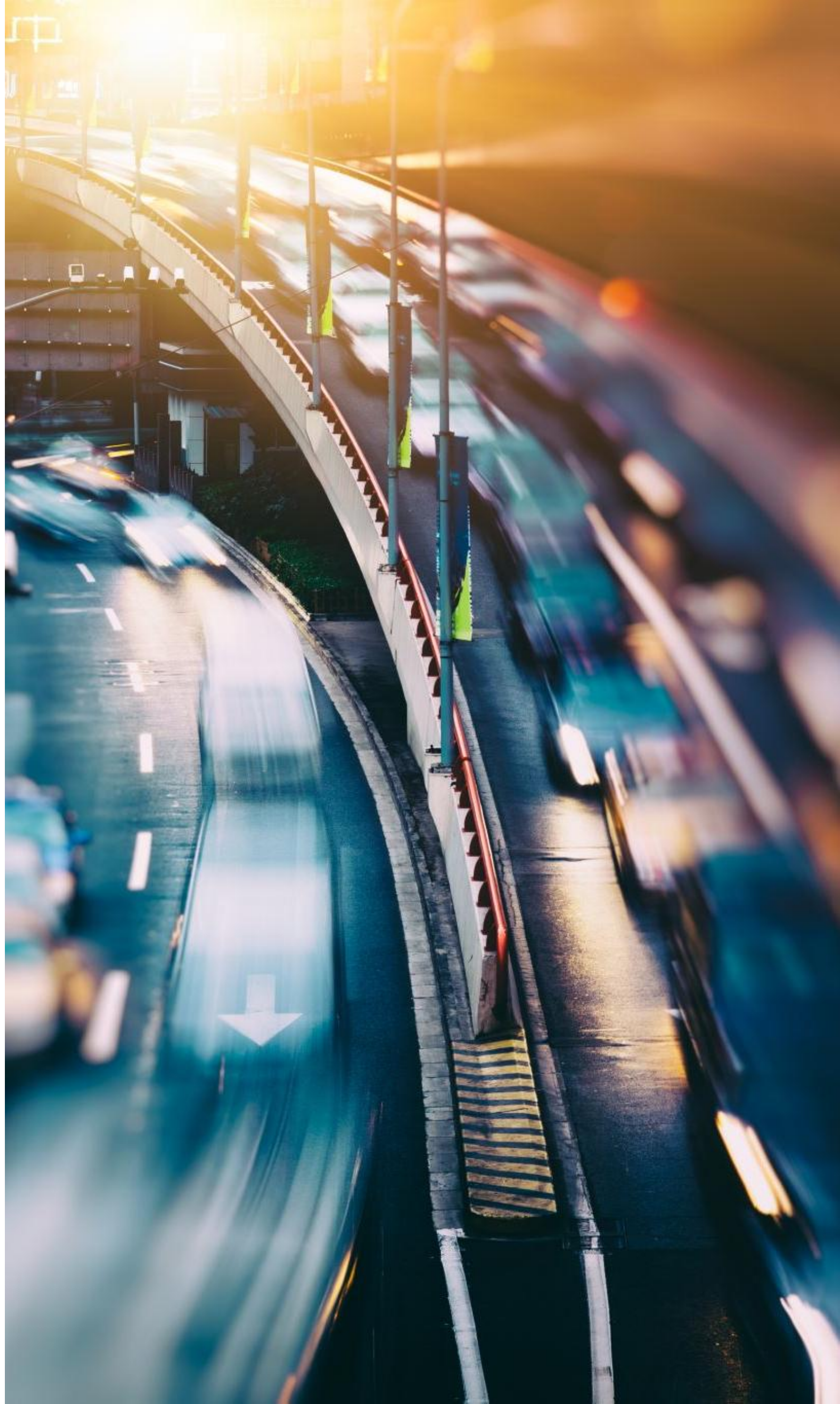
Operators can act **before** the customer leaves (offers, fixes, discounts)

📊 Optimize Network Resources

- **What it means:**
  ML learns how network resources are used and **allocates them intelligently**.
- **Examples:**
- Dynamic bandwidth allocation
- Smart load balancing between cells
- Power control optimization
- **Why it matters:**
- Better performance with **same infrastructure**
- Lower operational cost
- Essential for **5G / Massive MIMO**

📊 Forecast Traffic Demand

- **What it means:**
  ML predicts **when and where traffic will increase**.
- **Examples:**
- Busy hours
- Events (stadiums, holidays)
- Urban vs rural demand
- **Why it matters:**
- Prevent congestion
- Plan capacity upgrades
- Improve user experience proactively



🚨 Detect Fraud

What it means:
ML detects **abnormal behavior** in real time.

Examples:

- SIM fraud
- Subscription fraud
- International call abuse

Why it matters:

- Traditional rule-based systems miss new fraud patterns
- ML adapts to **new attack strategies**

⚙️ Fix Network Issues Faster

What it means:
ML identifies faults **before users complain.**

Examples:

- Predict cell failure
- Detect degrading KPIs
- Root-cause analysis

Why it matters:

- Reduced downtime
- Faster maintenance
- Lower OPEX

😊 Improve Customer Experience (QoE)

What it means:
ML links **network KPIs → real user satisfaction.**

Examples:

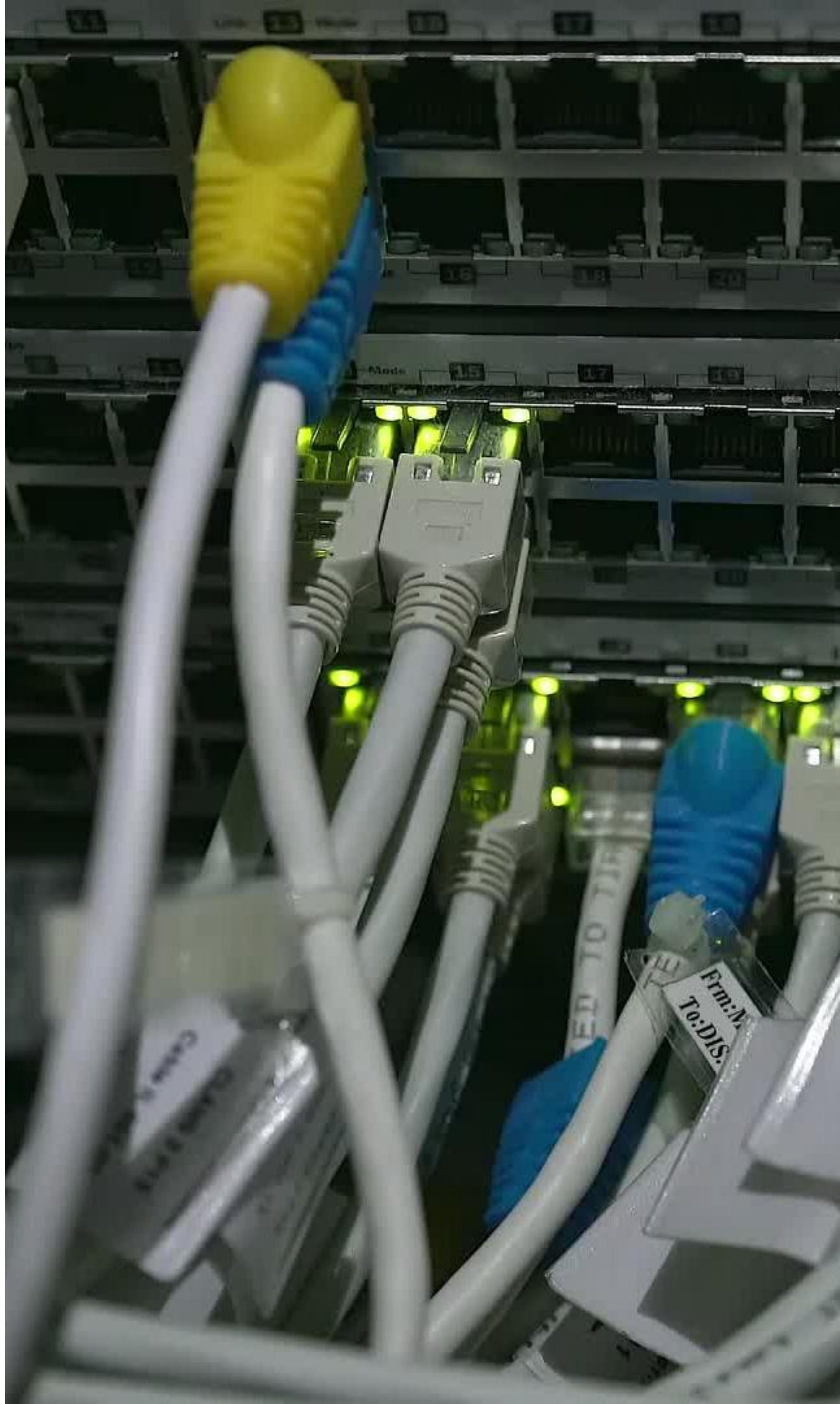
- Video buffering prediction
- Call quality perception
- App-level performance

Why it matters:

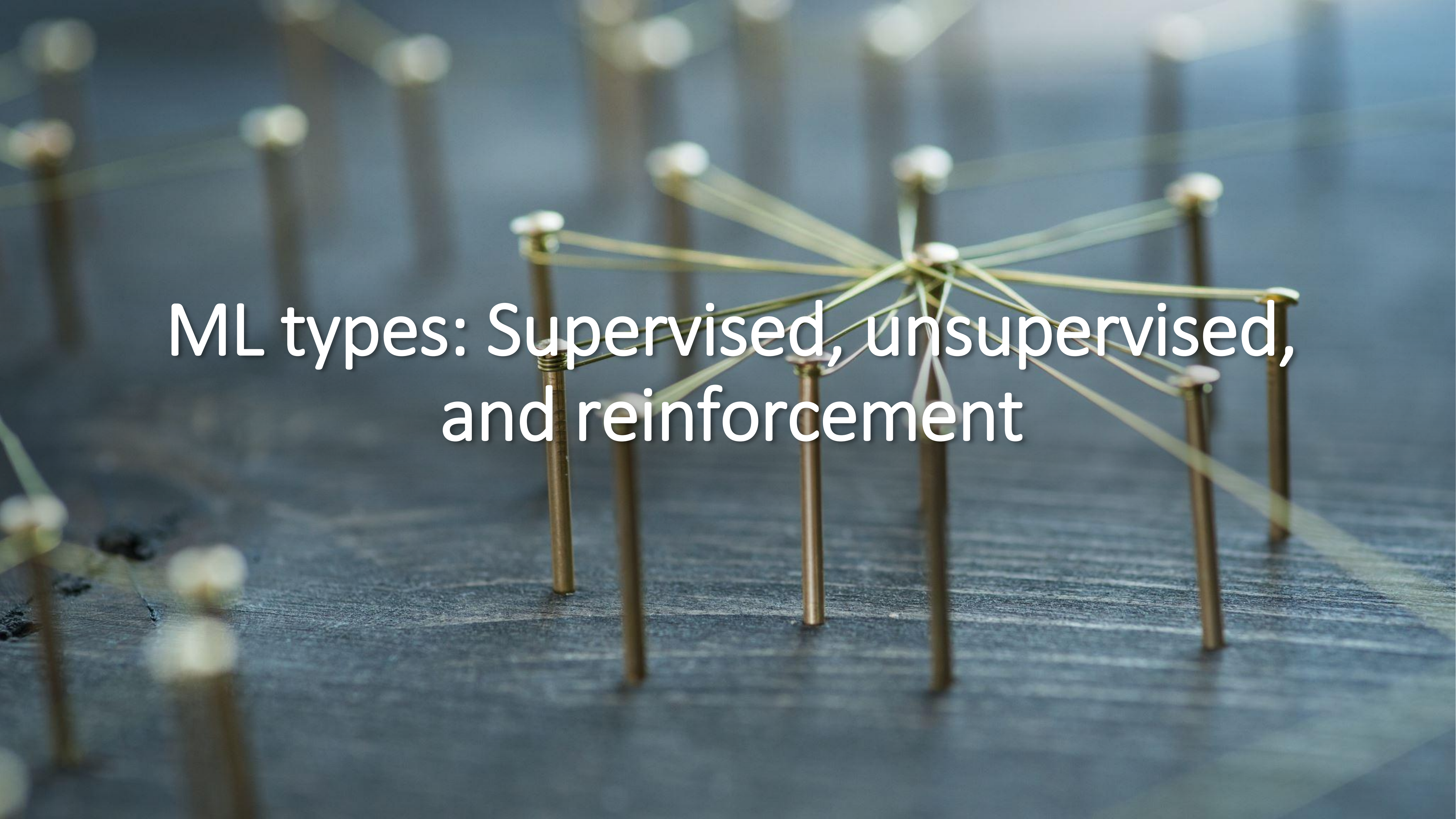
- Users care about experience, not KPIs
- Enables **QoE-driven optimization**

Summery

Machine Learning turns telecom networks from reactive systems into predictive, intelligent networks.



ML types: Supervised, unsupervised, and reinforcement



Supervised Learning (Most Common in Telecom)

🔹 **Telecom Example**

- **Input:**
- Signal quality (SINR, RSRP, RSRQ)
- **Output (label):**
- Throughput / data rate
- The model learns the **relationship** between:

Signal quality → Throughput (Goodput ≤ Throughput ≤ Bandwidth)

Unsupervised Learning (Pattern Discovery)

🔹 **What it is**

- Only **input data**
- No labels, no predefined answers
- Model finds **hidden patterns**

🔹 **Telecom Example**

- Input: signal quality data only
- No throughput, no "good/bad" labels
- Model automatically groups data into:
- Good signal areas
- Poor signal areas
- This process is called **clustering**

🔹 **What It Helps With**

- Coverage analysis
- Detecting weak zones
- Understanding network behavior
- Preprocessing data before supervised ML

⚠️ Not used for prediction, but for **understanding**

Reinforcement Learning (RL) (Learning by Experience)

🔹 **What it is**

- Learning via **trial and error**
- Model takes actions
- Gets **reward** (good decision) or **penalty** (bad decision)
- Think:
- Training a dog with treats 🐶
  Treat = reward, mistake = penalty

🔹 **Telecom Example**

- Model adjusts resource allocation
- If throughput improves → reward
- If performance degrades → penalty
- Over time:
- Learns which actions maximize performance

🔹 **Typical Use Cases**

- Dynamic resource allocation
- Power control
- Handover optimization
- Self-organizing networks (SON)
- ⚠️ Complex, needs careful design
  ⚠️ Less common than supervised learning in production

Goal	ML Type
Predict known outcomes	**Supervised**
Discover patterns	**Unsupervised**
Learn optimal actions	**Reinforcement**

🎯 How to Choose the Right ML Type

- 🔑 **One-Line Summary**
- **Supervised** → Predict
- **Unsupervised** → Understand
- **Reinforcement** → Optimize by experience

📡 Key Telecom Insight

Supervised learning dominates telecom because we usually have:

- Historical data
- Known inputs (signal quality)
- Known outputs (throughput, QoS, churn)

Supervised learning: Learning from labeled data

What is Supervised Learning?

- Supervised learning is a type of machine learning where:
  - You provide **input data**
  - You also provide the **correct output (label)**
  - The model learns the relationship between them

👉 If you **know what you want to predict**, supervised learning is the right choice.

Example 1: Throughput Prediction in a 4G Network (Regression)

📡 **Inputs and Outputs**

- **Input (x):** SINR (Signal-to-Interference-and-Noise Ratio)
- **Output (y):** Throughput (data speed)
- 🧠 **What the Model Learns**
- Higher SINR usually means higher throughput
- The model learns a mathematical relationship:
- Throughput=f(SINR)

🔮 **Prediction**

- New SINR value → model predicts the expected throughput

🔹 **Why This Is a *Regression Problem***

- Output is **continuous** (Mbps)
- It changes smoothly with the input

Linear vs Nonlinear Relationships

- With **one input (SINR)** → relationship may look linear
- With **multiple inputs**, such as:
  - SINR
  - Signal strength
  - Modulation & coding
  - Interference
  - Cell load
- ➡️ The relationship becomes **nonlinear**
- ✅ Supervised learning can handle **both linear and complex nonlinear patterns**

Example 2: Customer Churn Prediction (Classification)

- **Goal**
- Predict whether a customer will:
  - **Stay**
  - **Leave**
- **Input Features**
  - Dropped calls
  - Average data speed
  - Monthly bill
  - Usage behavior
- Output
  - Binary level:
  - 0 → Stay
  - 1 → Leave

Why It Matters

- Identify high-risk users early
- Take action before customers churn

Other Supervised Learning Use Cases in Telecom

📶 Network Performance Prediction

  - Predict user speeds in new areas or deployments
- Support realistic service promises and pricing

⚙️ Network Issue Detection

- Classify sites as:
  - Normal
  - Degraded
  - Faulty
- Engineers fix problems faster

📊 Traffic Forecasting (Time Series)

- Input: past traffic + timestamps
- Output: future traffic levels
- Used for:
- Capacity planning
- Congestion prevention
- 📌 Still supervised learning because **past outputs are known**

Unsupervised learning in Telecommunications

What Is Unsupervised Learning?

Unsupervised learning means:

- You only provide **raw data**
- There are **no labels** (no correct answers)
- The model **explores the data on its own**

👉 The goal is not prediction, but **discovery**:

- Patterns
- Groups (clusters)
- Unusual behavior (anomalies)
- ***Key Idea: We don't know the outcome in advance.*** *So, We let the machine tell us what the data looks like.*

Example 1: Customer Usage Behavior (Clustering)

📊 Input Data

- Voice usage
- Data usage
- Call duration
- Session frequency

🧠 What the Model Does

- It groups customers into **clusters**, such as:
- Heavy data users
- Heavy voice users
- Balanced users
- Low-usage users

💡 Why This Matters

- **Marketing**: Offer data-heavy plans to data users
- **Engineering**: Add capacity where data users dominate
- **Business**: Personalised services instead of one-size-fits-all

Example 2: Interference Zone Identification

Telecom Metrics

- **RSSI** (Received Signal Strength Indicator)
- **RTWP** (Received Total Wideband Power – interference indicator).

What the Model Does

- By feeding RSSI / RTWP data from many sites:
- Groups locations into:
  - Low interference
  - Medium interference
  - High interference zones.

Engineering Benefit

- Engineers focus on **worst interference areas**
- Faster troubleshooting
- Better QoS improvement

Example 3: Customer Segmentation for Planning

- **Unsupervised learning can help decide:**
  Where to expand capacity
- **What type of capacity is needed (voice vs data)**

Example:

City with many heavy data users → add data capacity

Area with many voice users → optimize voice resources

Example 4: Anomaly Detection (Unusual Behavior)

🚨 **What Is an Anomaly?**

A sudden deviation from normal behavior.

📉 **Telecom Examples**

Call drops:

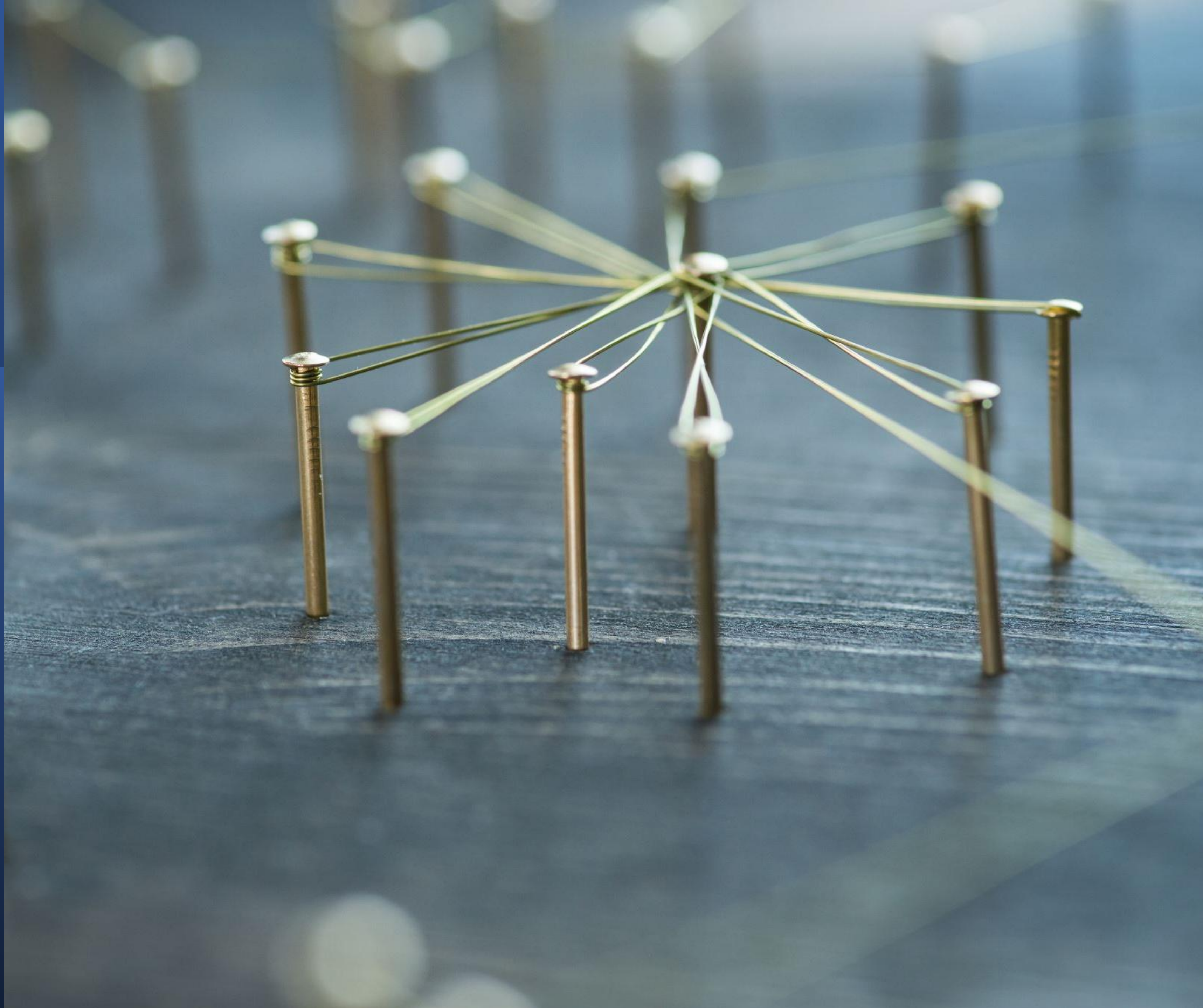
- Normal: 2 per day
- Suddenly: 50 → 🚨 anomaly

City-wide drops:

- Normal: 100 per day
- Suddenly: 2000 → 🚨 serious issue

- 🛠️ **Why It's Critical**
- Early detection of:
  - Network faults
  - External interference
  - Outages
- Faster response before customers complain

Reinforcement Learning



What Is Reinforcement Learning?

- Reinforcement Learning is a type of machine learning where:
- The system **learns by trial and error**
- It takes an **action**
- Receives **feedback**:
  - ✅ Reward (good decision)
  - ❌ Penalty (bad decision)
- Over time, it learns the **best actions**
- 👉 Think of it as:
- Learning by experience, not by examples.

Core Idea (Very Important)

No fixed labels. No predefined answers.
The model learns by interacting with the environment.

Telecom Example: Power Control as a Feedback Loop

📡 **Scenario**

- A user connects to a base station:
- User device transmits power
- Base station monitors signal quality.

⚖️ **The Trade-off**

- Too much power → ❌ interference
- Too little power → ❌ call drop
- Just right → ✅ stable communication
- This continuous adjustment forms a **feedback loop**.

Reinforcement Learning Solution

- The base station **tries an action**:
  - Increase power
  - Decrease power
- Then it checks the result:
  - Good signal → ✅ reward
  - Bad signal → ❌ penalty
- Over time:
  - The system **learns the best power level**
  - Power is adjusted automatically
- **Summary:**
- Reinforcement learning learns by trial and error to keep the connection stable.

Dynamic Resource Allocation

- What the network does:
- **Adjust bandwidth**
  - Give more bandwidth to busy users
  - Give less bandwidth when traffic is low
- **Change modulation**
  - Good signal → use higher modulation (faster speed)
  - Poor signal → use lower modulation (more reliable)
- **React to traffic load**
  - Many users → share resources smartly
  - Few users → reduce unnecessary resource usage
- **Result:**
- Better speeds
- Less congestion
- Efficient use of spectrum

Power Control Optimization

- What the network does:
- **Keep signal strong**
  - Enough power to avoid call drops
- **Reduce interference**
  - Avoid sending too much power
  - Protect other users
- **Save energy**
  - Use only the needed power
  - Reduce battery usage and network energy cost
- **Result:**
- Stable calls and data
- Cleaner radio environment
- Longer battery life

Case Study: How telecom networks use ML to optimize user throughput

Hayder AL-Salihi

Why Internet Speed Changes

- Sometimes internet is fast
  - Sometimes it is slow
  - Even when signal looks strong
- **Reason:**
  👉 Signal quality affects **throughput** (data speed)
- Machine Learning helps networks:
  - Understand this
  - Improve your experience automatically

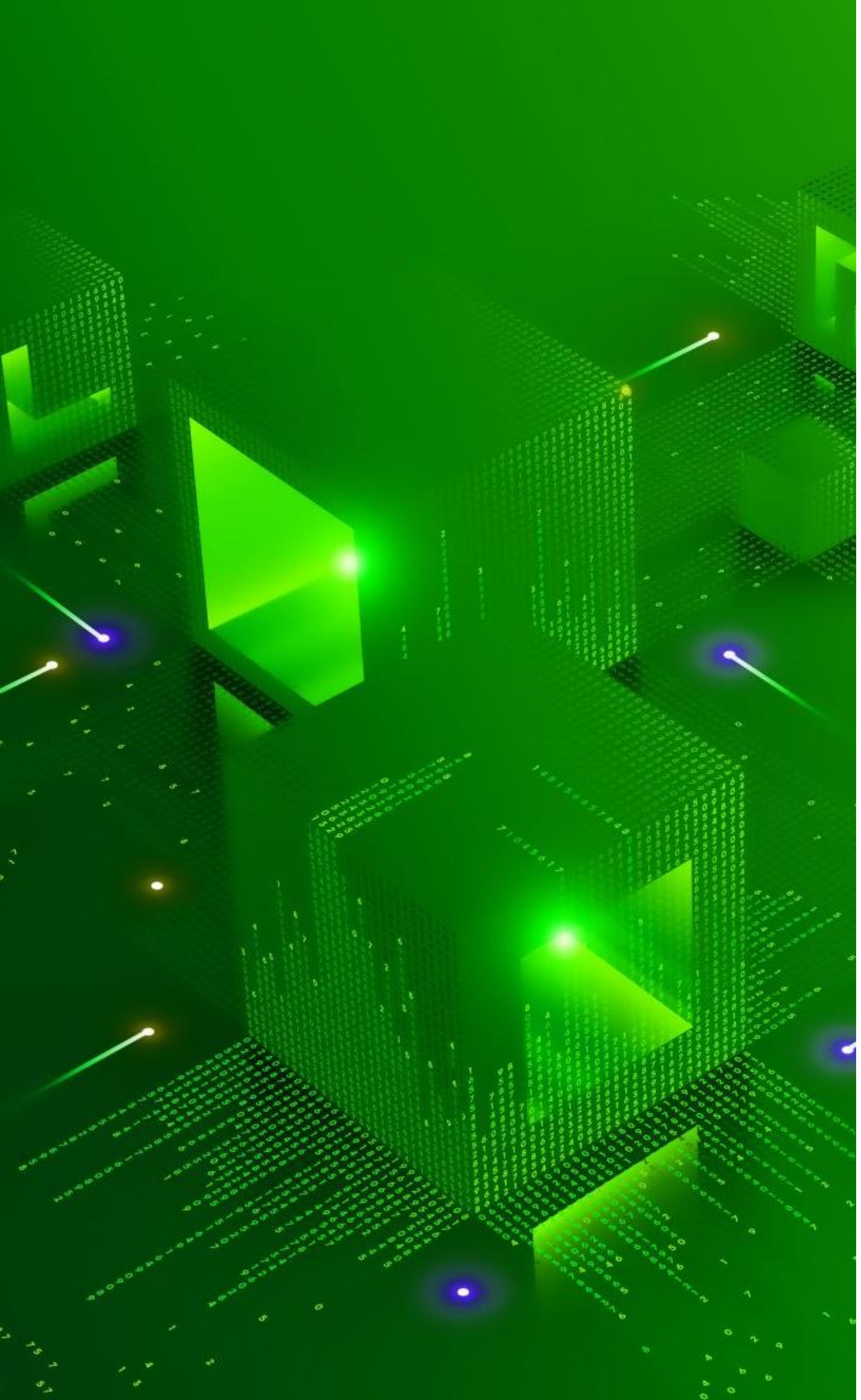
Data Collection

- Telecom networks **collect data all the time** from:
  - Thousands of cell towers
  - 4G and 5G networks
- They collect:
  - Signal quality (SNR, RSRQ)
  - Throughput (data speed)
  - Location
  - Time of day
- Data is stored in systems like:
  - NMS (Network Management System)
  - Intelligent controllers

Step 2 – Data Preparation

- Before using Machine Learning:
- Data must be **clean**
- The network removes:
  - Missing values
  - Wrong data
  - Outliers
- Inconsistent entries

✅ Only **high-quality data** is used
✅ Large machines (GPU servers) process the data



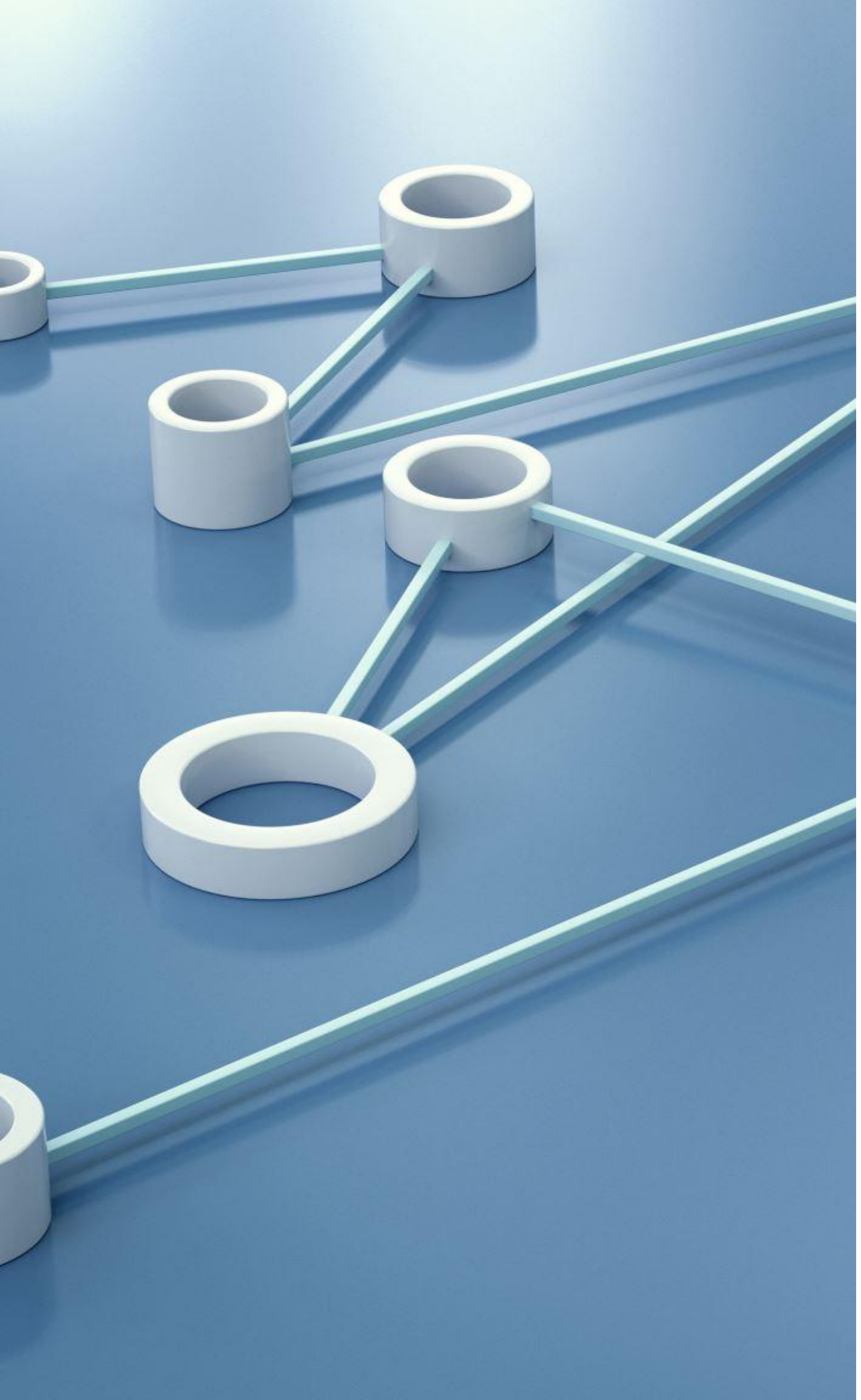
Model Training (Learning Patterns)

- The machine learns from data:
  - **Input:** Signal quality (SNR)
  - **Output:** Throughput (speed)
- It learns patterns like:
  - Good signal → high speed
  - Poor signal → low speed
- Other factors may also affect speed:
  - Congestion
  - Interference
  - Device type
- Machine builds a **predictive model**

Step 4 – Prediction

- Now the model is trained:
  - New signal quality value is given
  - Machine predicts expected speed
- Example:
  - Signal quality at your location = X
  - Predicted speed = 10 or 15 Mbps

👉 The network now **knows what speed to expect**



Step 5 – Network Actions

Based on predictions, the network can:

- Adjust settings
- Shift users between towers
- Add capacity where needed

⚙️ This happens **automatically.**
⚙️ No need to wait for human action.

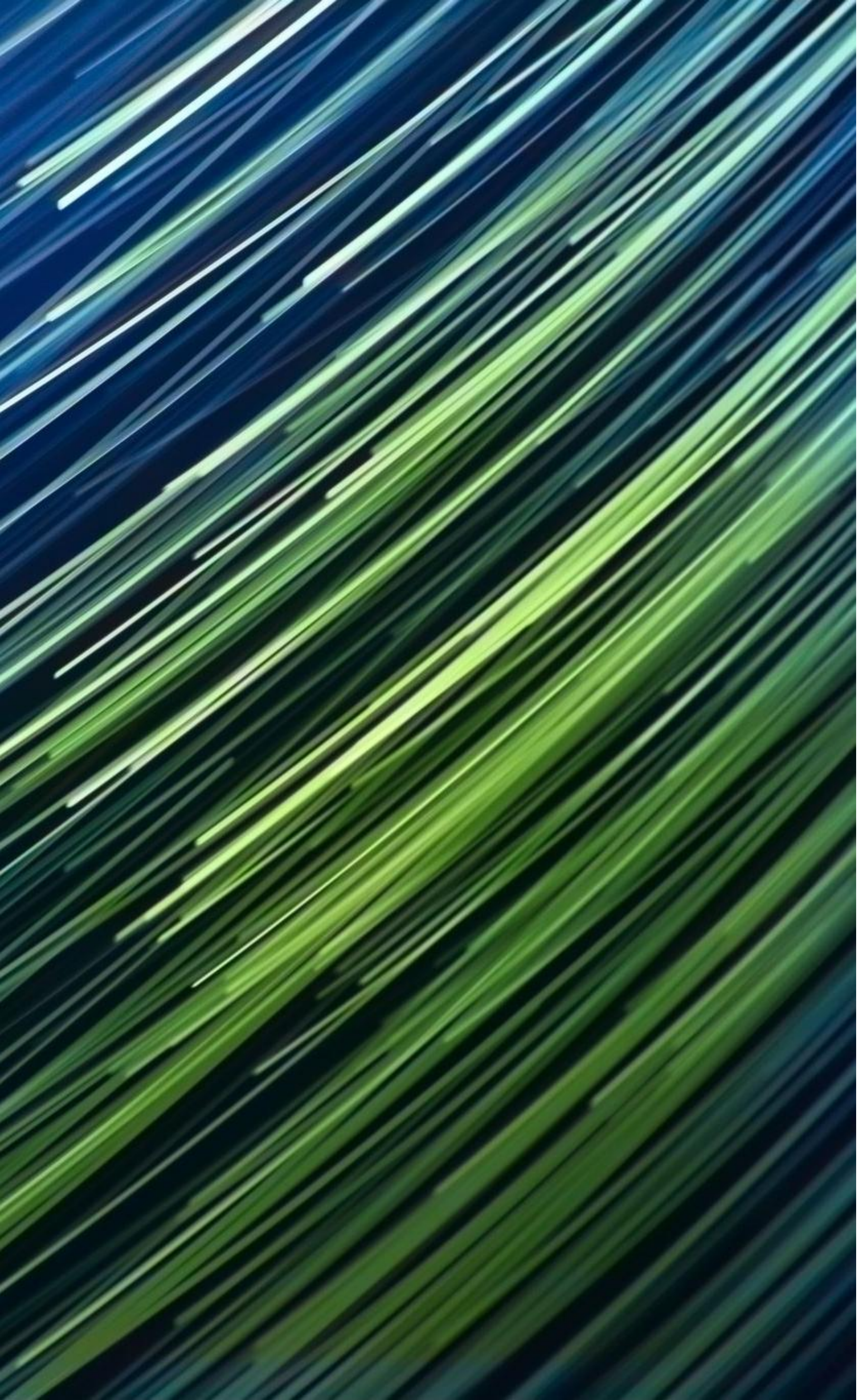
Why This Matters?

Machine Learning helps telecom networks to:

- React faster to problems
- Reduce downtime
- Optimize performance automatically

Improve user experience

- ⚠️ Important rule:
- Good data → Good predictions
  Bad data → Bad predictions



Final Message

When your internet is fast or slow, a machine behind the scenes is learning from data and trying to make your network better.

Machine Learning is transforming telecom networks one prediction at a time.