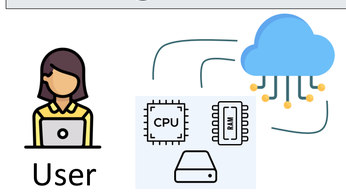
Is Machine Learning Necessary for Cloud Resource Usage Forecasting?

Georgia Christofidi Konstantinos Papaioannou Thaleia Dimitra Doudali

IMDEA Software Institute, Madrid, Spain

1. Problem Space

Challenge 1: Low resource efficiency in the Cloud.



User

Asks: X **> Y (Uses)** → **Resource Waste**

Accurate Forecasts can:
- ↑ Resource Efficiency
- ↓ Costs
- ↑ Energy Efficiency
- ↑ Application Performance

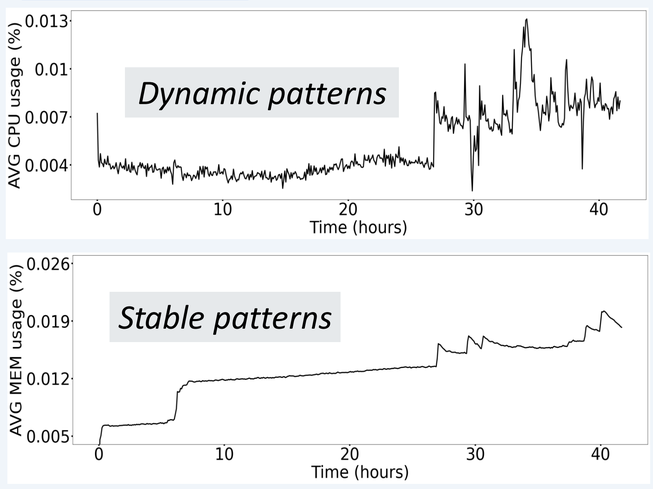
Approach: Future Resource Usage Forecasting.

Input: Past Resource Usage x1, x2, …, xn → **Forecasting Models** (ML, Statistical, Heuristic, Hybrid) → *Output: Future Resource Usage* $X_{n+1}, X_{n+2}, \ldots, X_{n+k}$

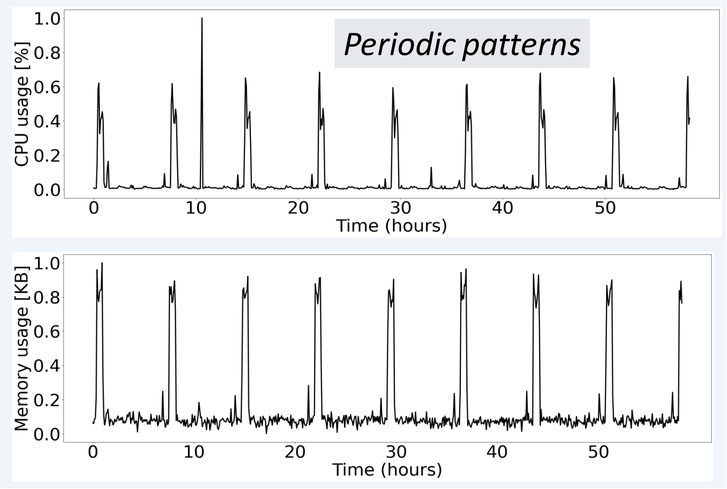
Problem: Achieving High Accuracy in Forecasting.

Challenge 2: Different Resource Usage Patterns.

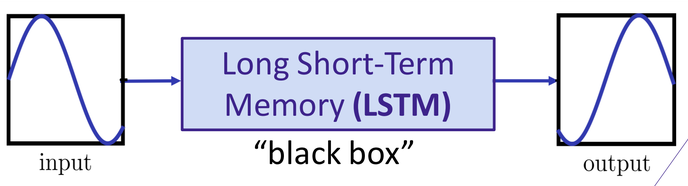
Workload level



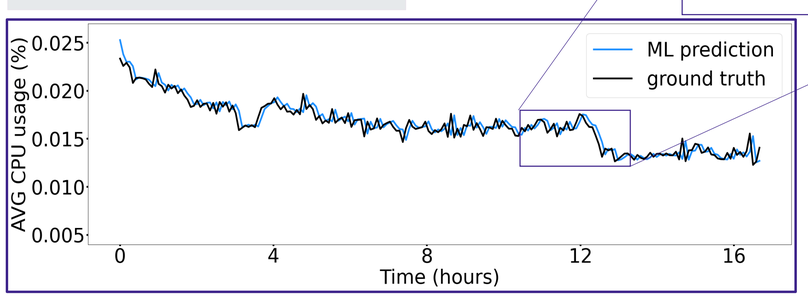
Virtual Machine level



2. Existing Approach

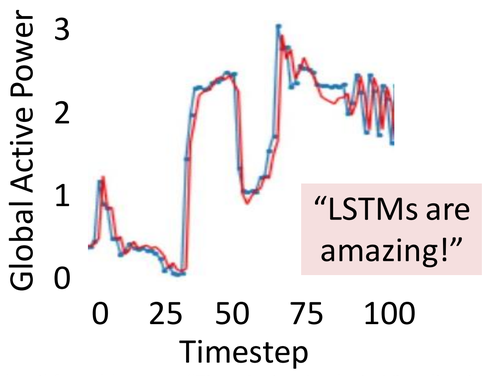


Usecase: Cloud Workloads



Our Insight: LSTM predictions resemble the **previous** timestep of the timeseries.

Usecase: Global Active Power Consumption



Source: Figure 12 detail from blog post "Time Series Analysis, Visualization & Forecasting with LSTM" on https://towardsdatascience.com

? Do we need ML to produce such **"shifted"** predictions?

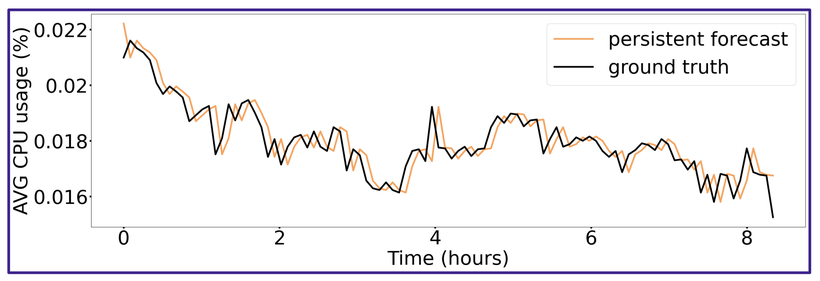
3. Proposed Approach

Persistent Forecast

Predict the value at the previous timestep.

For each value y(t) in the timeseries:
predict(t) = y(t - 1)

- ✓ Simple
- ✓ Lightweight
- ✓ Application agnostic
- ✓ No overheads



4. Experimental Results

Experiments with datasets across different:

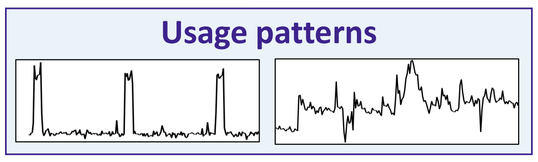
Cloud providers: Microsoft Azure, Bitbrain, Google Cloud, Alibaba Cloud

Resource Levels: Physical/ Virtual Machine, Workload

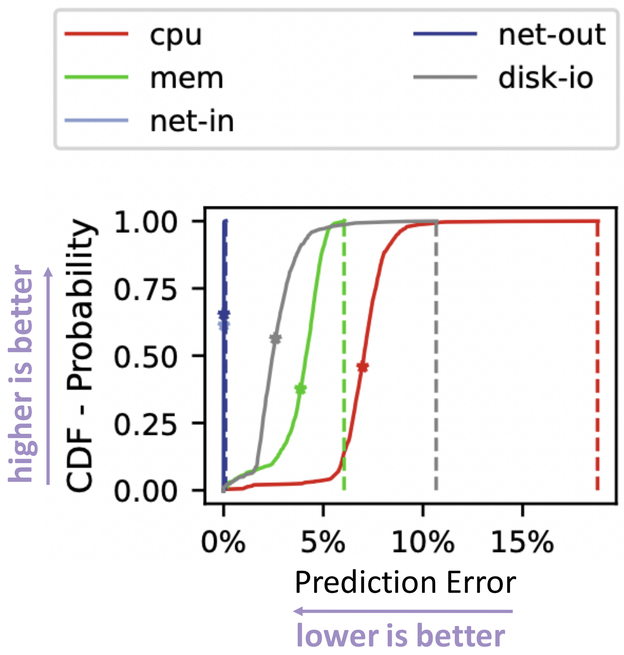
Resource Types: CPU, RAM

Frequency: Hourly/ Daily/ Weekly Windows

Usage patterns



All code is open source and available on Github.

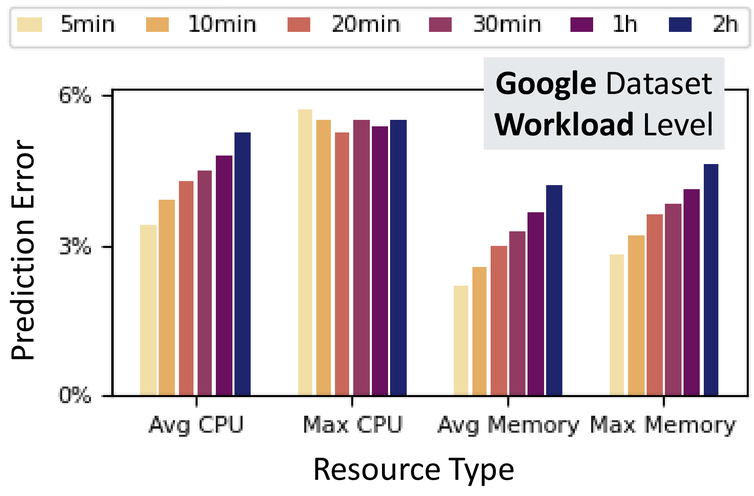


Alibaba Dataset
Physical Machine Level

We observe **very low error** values that depend on the resource type.

Takeaway: Persistent Forecast is **highly accurate for cloud data**, across resource types, levels of use and measurements.

? Sensitivity to the length of the time window



Takeaway: Small sensitivity to the **time window.** Opportunity for low error when window and patterns align.

Our Insight: The persistent forecast is **effective** because resource usage values of cloud workloads and servers, **persist over time**.

5. Summary

Machine Learning is **not always** necessary for Cloud Resource Usage forecasting.

Open Questions

1. When to use ML?
- exact use case
- data pattern

→ predictions ⇅ system's performance and decision-making

2. Which ML method to use, *when necessary*?

Probably not LSTMs

Other state-of-the-art ML methods for timeseries forecasting

Suggestions

1. Revisit existing systems and study the **data patterns**.

Values persist over time? → Try the **Persistent Forecast.**

2. Insightful and **judicious** use of ML, simple mechanisms to the extent possible.

Scan for code and paper: