Report date: 10/02/2025

Paper report:

**Overview:**

The paper is very confusing and I am tempted to keep it out of the literature study. It takes into consideration a lot of different techniques without explaining them properly and without providing context on why they could be useful in the machine unlearning field.

The first and second chapters are good, they introduce the issue of machine unlearning, the mention the first contributions to the field and they report a classic definition of machine unlearning (not mathematical).

They then proceed in listing MU techniques in 3 different subchapters:

* Data Deletion
  + Data poisoning
  + Data subsampling
  + Data shuffling
* Data Perturbation
  + Data anonymization
  + Differential privacy
  + Inverse data generation
* Model Update
  + Regularization
  + Transfer learning
  + Model pruning
  + Distillation
  + Model inversion

Most of these techniques are not deeply related to MU, they could be used and have ties with MU but the paper does not explain to what extend and doesn’t explain their employment in this field.

The 4th chapter provides a useful table listing many different datasets used in this field.

5th chapter explain challenges and potential solutions, they are very vague and provide challenges not strictly related to MU and very general:

* Attack sophistication
* Lack of standardization
* Lack of transferability
* Lack of interpretability
* Lack of training data
* Resource constraint

Overall the paper is very vast but it doesn’t really say anything new or relevant, many general issues, challenges and techniques used in AI are reported without saying what they have to do with MU.

I’m clarifying with Sebastian and Stefan if I’ve missed anything in this paper of if the paper itself is actually not very useful.

**Authors:**

Thanveer Shaik et alii.

**Year:**

2024