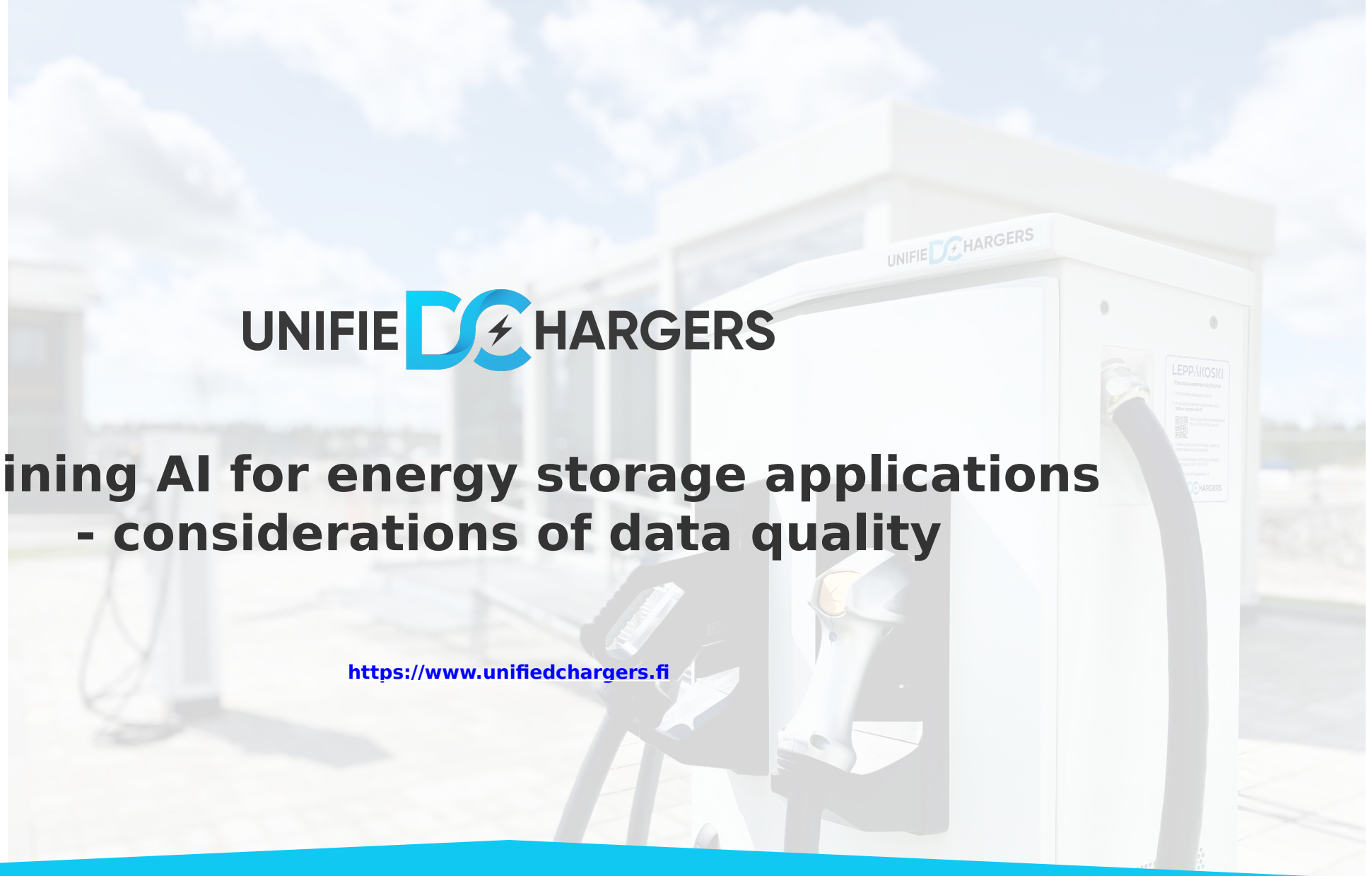


Training AI for energy storage applications - considerations of data quality

https://www.unifiedchargers.fi

Jukka Anttonen, CEO/CTO
+358 40 765 5038
jukka@unifiedchargers.fi

How to train AI? 4 step program!

Step 1: Dataset

Step 2: Validate

Step 3: Test

Step 4: Repeat

A few basics

- Computers and algorithms are very good at performing well defined tasks
- Tasks are defined by humans
- AI is (for now at least) unable to define the tasks for it self
- AI/computer can perform a continuous task based on some initial values and iterate from that based on new data generated or use other statistical methods
- AI can be thought to be a statistics analyzing tool (in most cases a powerful one)
- Before training there have to be some initial dataset to start the training proses
  - OR not...

Old news but an excellent example

- well defined task
- no initial dataset (except rules of the game)
- pure iterative approach

AlphaZero AI beats champion chess program after teaching itself in four hours

Google's artificial intelligence sibling DeepMind repurposes Go-playing AI to conquer chess and shogi without aid of human knowledge

▲ AlphaZero's victory is just the latest in a series of computer triumphs over human players since Computer programs have been able to beat the best IBM's Deep Blue defeated Garry Kasparov in 1997. Photograph: 18percentgrey / Alamy/Alamy

AlphaZero, the game-playing AI created by Google sibling DeepMind, has beaten the world's best chess-playing computer program, having taught itself how to play in under four hours.

-The Guardian 2017

What about energy storage system?

Problem: What state our energy storage should be at any given time?

- Energy storage systems cannot be represented with purely iterative models like Google approached the game of chess.

- There needs to be some historical or simulated(a.k.a educated guess) dataset and the proses turns more into a statistical analysis.

Data quality assessments are needed when dealing with any not finite dataset. Preferably done by humans.

“Data quality assessment (DQA) is the process of scientifically and statistically evaluating data in order to determine whether they meet the quality required for projects or business processes and are of the right type and quantity to be able to actually support their intended use.”

-Technopedia

Classic data quality issues and some of their results

- Lack of data over the whole “area” or data points are too similar with each other
  - leads to self iteration
- Data points are given too high or too low weight based on age or source etc.
  - leads to self iteration
- New data points are generated too fast or too slow by the program (noise)
  - leads to overreacting or under-reacting
- Too much data
  - systems in overloaded and cannot make predictions in “real-time” or fast enough
- Dataset doesn’t represent the phenomena or the source is not high enough quality
  -unpredicted behavior
- Correlation is not necessarily a causation!

Classic data quality issue example

Lack of data over the whole “area” or data points are too similar with each other:

Estimating a remaining battery capacity in ESS can be problematic when using only statistical methods

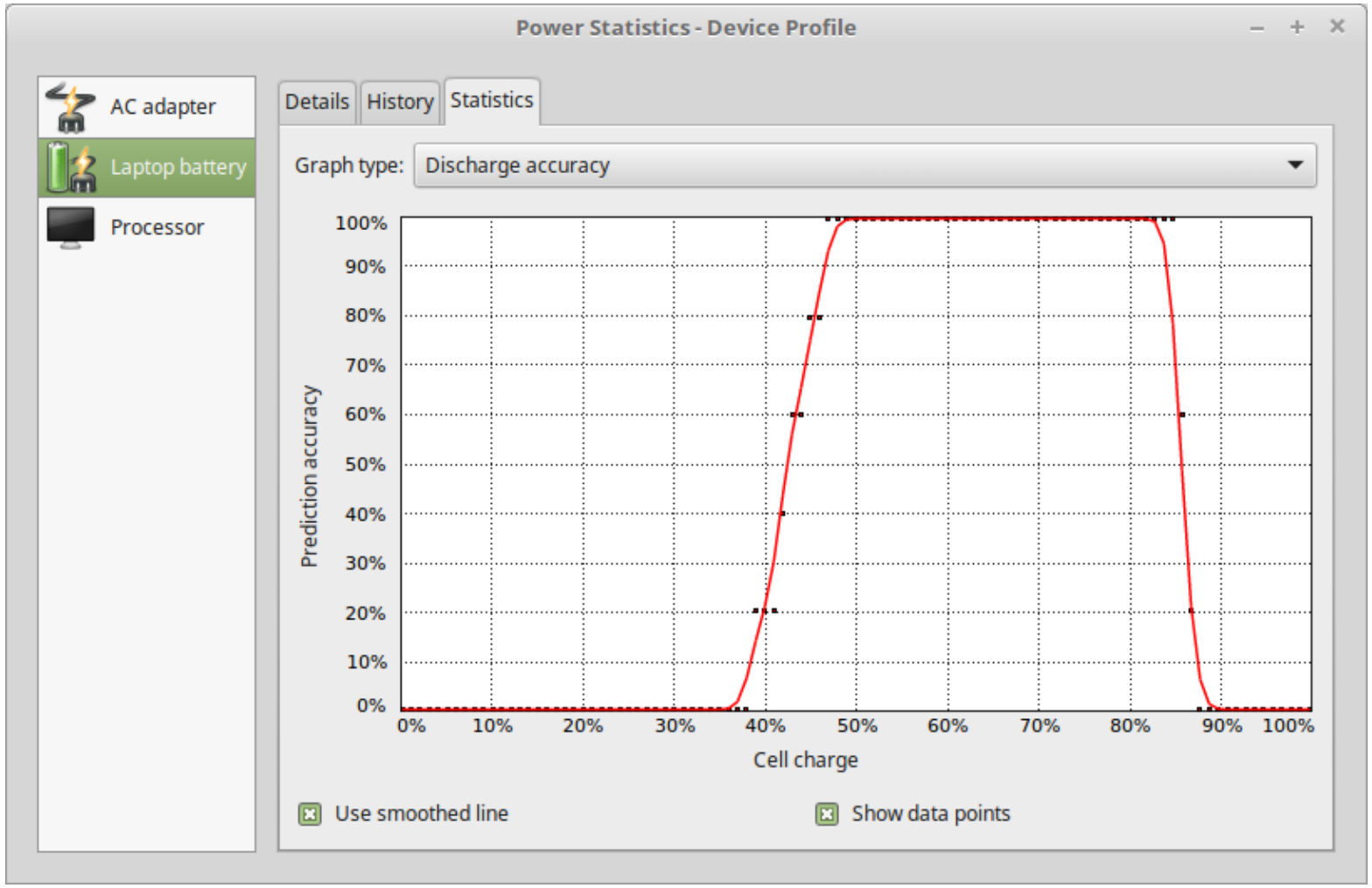
- Dataset usually has a very few data points in low and high states of charge because those are unwanted states of the battery
- Typical use happens between 30-80% state of charge most of the time in ESS application
- Data points in that region gets over represented unless weight down separately
- System might “self iterate” even though the data set would be carefully weight

- Result: False battery degradation

80% comes a new 100% and 30% comes a new 0% -> large portion of the battery capacity remains unused.

Another example of this is a laptop battery

- I rarely use all of my laptop battery
- This particular ”AI” says it can be 100% accurate in estimating remaining battery power between 50 – 80% but it has no idea how my battery will behave outside of that range because there is very little data outside that range.
- I have cheated this particular machine learning software accidentally by setting a limit for charging only upto 80% thus preventing the battery being fully charged to 100%. Apparrently the software is unable to check system settings.
- **Solution** would be simply have more datapoints over larger area or give less weight on values that are similar to values next to it. Preferably both and access to critical paramers of the system.



Disclamers

- Both examples used previously are out of proportions and out of context
- Generally all these issues and phenomena are taken into consideration already
- In real-life applications the models and data sets are more sophisticated than illustrated here

- But it is always worth asking what is the problem we are solving?
- Well defined problems and tasks will produce the best outcome

Business side of AI and EV-charging

- EVs are basically energy storages on wheels. Untapped resource mostly today.
- Better electrical grid stability when charging is controlled “wisely”
- Better electrical grid stability leads to better energy balance in the grid and lower emissions in energy production
- So far lower emissions in energy production have also predicted lower energy prices. Could be even more than just a correlation!
- By developing new and training further existing models there is a large emission reduction potential. But to fully utilize it we need to better utilize the end users input.

  -End user input here being simple as:
  “When do you think you are going to unplug your car next time?”
- Predicting future is hard but sometimes it useful to just ask from someone who knows.

Q&A