Consideration of the Real World in Use-Case-Based Benchmarking

Kay Bierzynski
02.07.2021

Motivation

- Academic data sets are mostly:
  - Perfect environmental conditions
  - Limited time frames
  - Limited situations
  - Outliers removed
  - Data is complete

- Real world:
  - Environmental conditions are not perfect
  - Long run times
  - Outlier can matter
  - Missing or not synchronized data
  - Unexpected Situations

Adverse and Changing Environmental Conditions

- Changing conditions:
  - Lighting
  - Day and night cycle
  - Temperature
  - Noise levels
- Weather:
  - Rain, snow, fog
  - (Thunder)storms, tornados
- Other:
  - Construction sites, factories, blast furnaces and more

Long Run Times

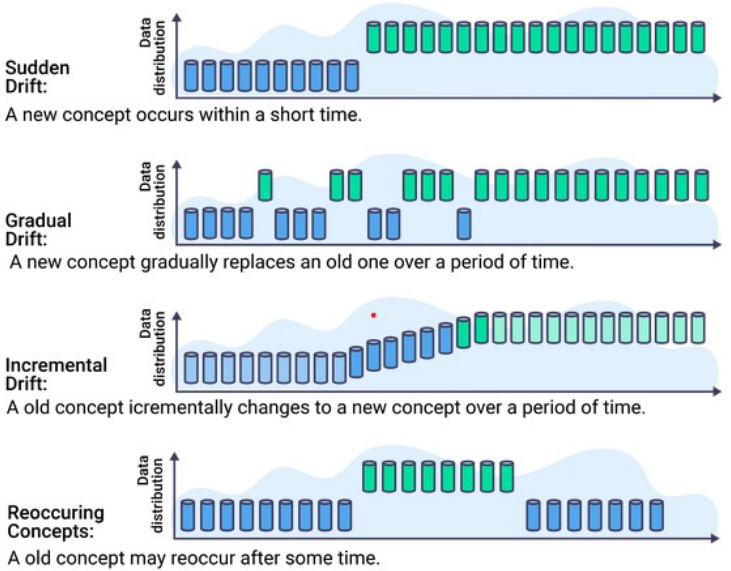
- Real world systems are planned to run for a long time in most contexts.
- This entails many challenges that need to be considered:
  - Seasons
  - Change of the environment e.g. construction of new buildings
  - User changing their behavior due to ageing, children and more
  - Ageing of hardware

- These challenges lead to **drifts**

Types of Drifts

- **Concept drift** – a change in the distribution of p(Y |X), meaning that there was a change in the relationship between the input of the model and the true label.
- **Prediction drift** – a change in the distribution of the predicted label – p(ŷ |X), meaning that there was a change in the relationship between the input of the model and the model's prediction.
- **Label drift** – a change in the probability of a label p(Y).
- **Feature drift** – a change in the probability of p(X), meaning there was a change in the distribution of the model's input.

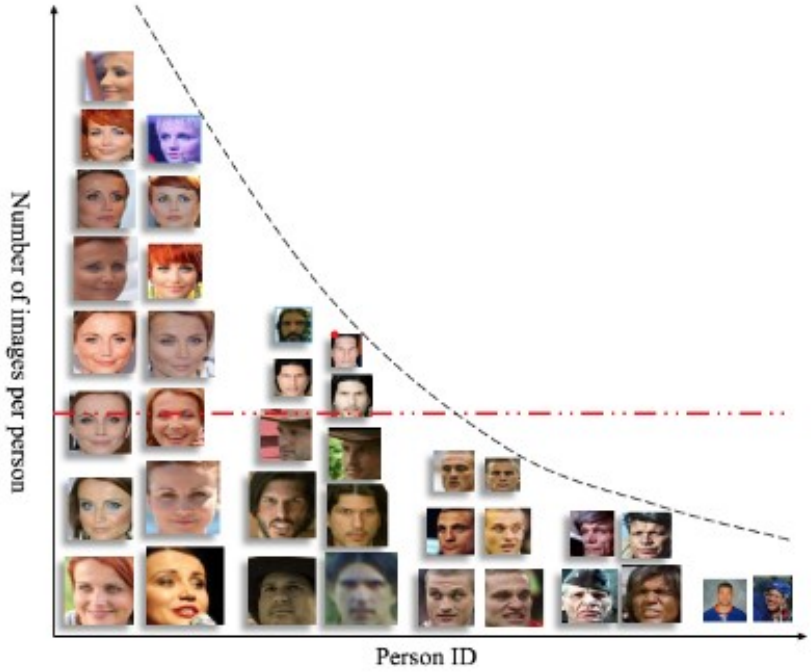
- Parameters:
  - X - Model's input population
  - ŷ - Model's prediction
  - Y - True label population.



[1]

The Importance of Outliers

- Real world data sets have often an head and tail distribution in regards to their classes.
- Ignoring outliers or classes with very few samples can lead to severe consequence e.g. car accidents or law suits due to perceived racism
- Solutions:
  - Train extra model for handling outliers
  - Transfer learning from head to tail



[2]

Missing or not synchronized data

- Not all data will always be available to the model for processing
- Reasons:
  - Sensor failure
  - Connection failure
  - Mobile sensors entering and leaving the network e.g. smart phones
  - Limiting update frequency because of energy management
- Depending on the application missing and not synchronized data should be considered as an test case

Unexpected Situations or Mobile vs Stationary Sensors

- It was always difficult to consider all situations a system can encounter during design phase
- This problem is becoming more severe due to:
  - Increasing reliance on AI
  - Emergence of more applications relying on sensors that move
- Moving sensors can encounter much more situations than stationary ones

Conclusion

- Academic data set are not considering some challenges encountered in a real world context due to reasons like data availability
- Test cases and benchmarking need to be carefully designed when researching and developing AI models and neuromorphic hardware that will be used for products

- How to improve benchmarking:
  - Simulate drifts via simulator tools or synthetic scripts
  - Exclude at least one class or situation from trainings data, but used in the test phase
  - Specify test cases with unsynchronized or missing data and evaluate the resulting behavior
  - Add noise to clean data samples
  - Consider for trainings and test data sets the time frame
  - Simulate the ageing of the hardware

Part of your life. Part of tomorrow.

References

- [1] https://www.aporia.com/concept_drift_in_machine_learning_101/
- [2]https://ai.tencent.com/ailab/media/publications/Range_Loss_for_Deep_Face_Recognition_with_Long-Tailed_Training_Data.pdf
- Other images are the property of Infineon Technologies AG