CSCE 590-1: From Data to Decisions with Open Data: A Practical Introduction to AI

Lecture 14: Time Analysis – Invited Speaker

PROF. BIPLAV SRIVASTAVA, AI INSTITUTE

2ND MAR, 2021

Carolinian Creed: "I will practice personal and academic integrity."

Organization of Lecture 14

- Introduction Segment
  - Recap of Lecture 13

- Main Segment
  - Invited talk

- Concluding Segment
  - About Next Lecture – Lecture 15
  - Ask me anything

Introduction Segment

Recap of Lecture 13

- Reviewed AutoAI paper
- Looked at time
  - Representation
  - Time difference analysis
  - Simple temporal analysis
  - Auto-regressive models
- A broad area for generating insights

Beyond Statistical Analysis of Time

- Events analysis
  - Event: Label + timestamp
  - Detecting and processing event
- Causal reasoning
  - Establishing cause-and-effect
  - **Not** the same as correlation

Main Segment

Lecture 14: Invited Lecture

- Understanding Timing Errors in Datasets
  - Sandeep Sandha, PhD Candidate, UCLA

    This lecture will introduce the different reasons for data timestamping errors present in the edge devices with a focus on Smartphones. We will discuss how the poor kernel design in Android is one of the main contributors to timestamping errors, with experiments showing the magnitude of error as high as 5 seconds across modern devices. To understand the impact of timestamping errors, the use-case of multimodal fusion for human activity recognition and cooking detection will be presented.
    Our analysis will highlight significant degradation in classifier accuracy due to the timing errors.

    As a solution approach to reducing the impact of timing errors on ML applications, we will introduce a new form of data augmentation called time-shift data augmentation. **As a motivational scenario, we will analyze the cooking dataset from the open-source domain.** Further, we will briefly look at the missing samples and the variable sampling rates in the data. Finally, we will conclude by presenting system techniques that can be deployed to enable robust data collection on the edge and introduce our open-source library used by UCLA's David Geffen School of Medicine to **timestamp data collected from the brain implants of patients**.

- **Keywords**: Sensor data, time analysis, AI on the edge

Related Papers and Codes

1. Time Awareness in Deep Learning-Based Multimodal Fusion Across Smartphone Platforms, IoTDI-2020
   1. Code: a) https://github.com/nesl/GoodClock
   2. https://github.com/nesl/CMActivities-DataSet
2. Deep Convolutional Bidirectional LSTM for Complex Activity Recognition with Missing Data. Human Activity Recognition Challenge - Smart Innovations, Systems and Technologies, Ch. 4, Springer Singapore (2020).
   1. Code: https://github.com/nesl/Robust-Deep-Learning-Pipeline
3. Exploiting Smartphone Peripherals for Precise Time Synchronization, ISPCS-2019
   1. Code: https://github.com/nesl/Time-Sync-Across-Smartphones
4. Enabling edge devices that learn from each other: Cross modal training for activity recognition. EdgeSys-2018
   1. Code: https://github.com/nesl/RecycleML

Lecture 14: Concluding Comments

- We heard about timing errors and their impact on devices
- Implications for their usage in detecting human activities and applications built using them

Concluding Segment

About Next Lecture – Lecture 15

Lecture 15: Reasoning

- All insights are not drawn from raw data
- Reasoning: drawing insights from knowledge (known or learnt)
- Various settings
  - Deterministic
  - Uncertain knowledge
  - Optimal decisions