Continual Stream Learning: A Comprehensive Framework for Continual Model Training and Evaluation in Streaming Environments with a Case Study*

TÜBİTAK Project No. 125E060

Fazlı Can (PI), Seyit Koçberber (I)

May 15, 2025

Topic. Traditional machine learning focuses on static datasets and struggles to adapt to changes in dynamic environments. Stream Learning (SL), on the other hand, aims to solve this problem by continuously updating the learning model online as new data becomes available. While it tracks current changes in data distribution (concept drift), it tends to ignore past information. Continual Learning (CL) tries to prevent catastrophic forgetting—loss of previously learned information—by using limited updates, but it doesn't address the concept drift issue. In this project, we introduce for the first time in the literature the Continual Stream Learning (CSL) model, which aims to overcome the limitations of the previous two approaches by handling changes in incoming data while simultaneously preventing catastrophic forgetting. Our work aims to develop a comprehensive framework for Continual Stream Learning (CSL) environments that can beadapted to different problems, and to demonstrate its use and effectiveness through a case study focused on classification.

Contributions. The unique contributions of our project are: Original Contribution (OC) 1: The formal definition of Continual Stream Learning (CSL) and the development of a framework for this problem area. OC 2: The development of a specialized neural network called the Short-Term Long-Term Memory Model (SLMM), designed to effectively prevent catastrophic forgetting while adapting to evolving data updates. OC 3: The creation of a new optimizer, Adaptimizer, designed to improve the model's adaptability to concept drift in continuous stream learning environments. OC 4: The introduction of a new metric, the Learn-Forget-Adapt Score (LFA Score), to evaluate the model's effectiveness in adapting to changing data and preventing catastrophic forgetting. OC 5: A case study demonstrating the effectiveness of our model.

Method. The core component of our project is the Short-Term Long-Term Memory Model (SLMM), a neural network designed to solve the catastrophic forgetting problem while adapting to changes in streaming data. Inspired by the human brain's memory system, this model consists of three main components: Sensory Memory, which quickly and effectively adapts to new data; Short-Term Memory, which stores data temporarily; and Long-Term Memory, a more complex structure that preserves

historical information to prevent forgetting. Sensory Memory transfers information to Short-Term Memory using a self-attention mechanism, while a transfer/retrieval gate controls the flow of information between Short-Term and Long-Term Memory.

Project Management. Our project consists of five work packages (WP), designed to achieve our original contributions: WP1: The creation of the NYT-Temporal and THD-Temporal datasets for use in a continual stream learning (CSL) environment. WP2: The development of the CSL framework, along with the necessary continuous training and testing functions. WP3: The implementation of our SLMM approach, which incorporates both short- and long-term memory units. WP4: The creation of the Adaptimizer, an optimization method designed to ensure the framework adapts effectively to temporal changes in the data. WP5: A case study that involves comparing the performance of our proposed model with baseline methods in classification tasks through comprehensive statistical experiments.

Impact. The Continual Stream Learning (CSL) approach, which we will define for the first time in the literature, will play a key role in application areas where it is important not only to learn new information but also to retain past knowledge. Examples include healthcare, cybersecurity, finance, monitoring court decisions, and recommendation systems. As always, the models we develop will be shared on GitHub by our research group. Our project will also benefit the Turkish research community by attracting domestic researchers to this topic through the contributions we make and the datasets we create. Furthermore, our project's scholars will advance in their careers, eventually training new researchers and connecting our work to the future.

* **Duration: May 15, 2025 - May 15, 2027**
RAs supported by the project: Pouya Ghahramanian, Sepehr Bakhshi
(Budget 1.677.900 TL / $43.351as of May 15, 2025: official project beginning date)