Anti-drifting Feature Selection via Deep Reinforcement Learning (Student Abstract) *

Aoran Wang[1], **Hongyang Yang**[2], **Feng Mao**[2], **Zongzhang Zhang**[1], **Yang Yu**[1], **Xiaoyang Liu**[3]

[1] National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
[2] Alibaba Group, Hangzhou 310052, China
[3] Columbia University, New York 10027, USA
{wangar, zhangzz, yuy}@lamda.nju.edu.cn, {yanghongyang.yhy, maofeng.mf}@alibaba-inc.com, xl2427@columbia.edu

Abstract

Feature selection (FS) is a crucial procedure in machine learning pipelines for its significant benefits in removing data redundancy and mitigating model overfitting. Since concept drift is a widespread phenomenon in streaming data and could severely affect model performance, effective FS on concept drifting data streams is imminent. However, existing state-of-the-art FS algorithms fail to adjust their selection strategy adaptively when the effective feature subset changes, making them unsuitable for drifting streams. In this paper, we propose a dynamic FS method that selects effective features on concept drifting data streams via deep reinforcement learning. Specifically, we present two novel designs: (i) a skip-mode reinforcement learning environment that shrinks action space size for high-dimensional FS tasks; (ii) a curiosity mechanism that generates intrinsic rewards to address the long-horizon exploration problem. The experiment results show that our proposed method outperforms other FS methods and can dynamically adapt to concept drifts.

Introduction

Feature selection (FS) is an indispensable procedure in data-driven machine learning pipelines. However, the data distribution could change dramatically as time goes, i.e., the concept drift phenomenon, usually causing the effective features change. Existing FS methods did not track the changing effective features when concept drifts happen. However, the reinforcement learning (RL) agent could capture the changes in data distribution by perceiving changed rewards during interactions with the environment and, eventually, rediscover the effective feature selecting strategy after concept drifting.

Selecting features with deep RL (DRL) is not a widely studied topic yet. Few related studies use a multi-agent framework (Liu et al. 2019). However, when the number of features scales up, training hundreds or even more FS agents through independent DQNs would face severe computation cost and non-stationarity, resulting in poor generalization.

*Corresponding author: Zongzhang Zhang. This work is supported by the NSFC (No. 62276126) and Alibaba Group through Alibaba Research Fellowship Program.

In this paper, we propose a general DRL framework to dynamically select effective features on data streams with concept drifts. Our framework adopts a single-agent design and addresses two key challenges. One is the problem of large discrete action space. We design a skip-mode environment to cut down the action space size by over 90%. Another major challenge is the problem of exploration with long horizon and sparse rewards. We address the problem via the intrinsic curiosity module (ICM) (Pathak et al. 2017), encouraging our FS agent to explore novel feature combinations.

Methodology

As shown in Fig. 1, our feature selector chooses the timely effective features by a trainable agent continuously interacting with the specialized skip-mode environment. In a single selection step, we first represent the present state s_t from the currently selected features I_t, and choose an action a_t by the RL agent's policy network $\mu(\cdot|\theta)$. By taking action a_t, we select the feature index specified by that action and obtain a new feature subset I_{t+1}. After taking action a_t, the reward evaluating model computes the performance metrics according to the selected data that corresponds to I_{t+1}, and returns the environment reward r_t^e. We further generate an intrinsic reward r_t^i by the curiosity mechanism to help exploration. Then, the trajectory $\{s_t, a_t, r_t^e + r_t^i, s_{t+1}\}$ is added to the replay buffer and $\mu(\cdot|\theta)$ is updated with the PPO algorithm.

A selection round starts from the empty feature set ($|I_0| = 0$) and repeats the above process until the termination condition is reached at the end of a complete feature selection. The RL feature selector runs the selection rounds along with the data stream to track timely effective features.

The Skip-mode Environment

As shown in Fig. 1, our proposed skip-mode environment consists of a discrete action space sized n, i.e., $A = \{a_1, a_2, \cdots, a_n\}$, when conducting FS on data with N features. By taking action a_i, where $i \in \{1, 2, \ldots, n-1\}$, the agent selects the next i-th unselected feature in the original feature sequence, skipping the $i-1$ features in between. By taking the special action a_n, the agent skips all of the next n unselected features. The selection round in one episode starts from the status in which no feature is selected, and

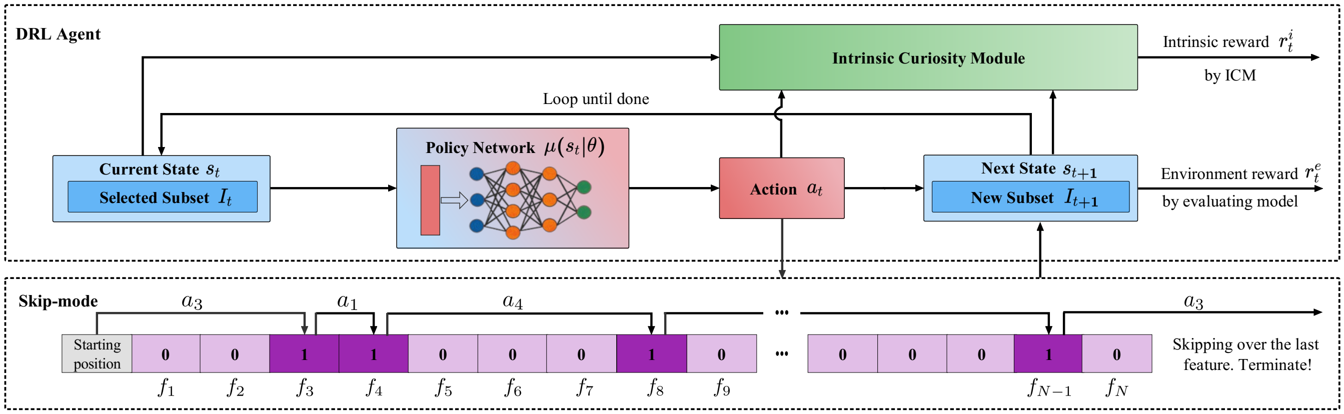

Figure 1: Our DRL feature selector, including the skip-mode environment, PPO-based agent, and intrinsic curiosity module.

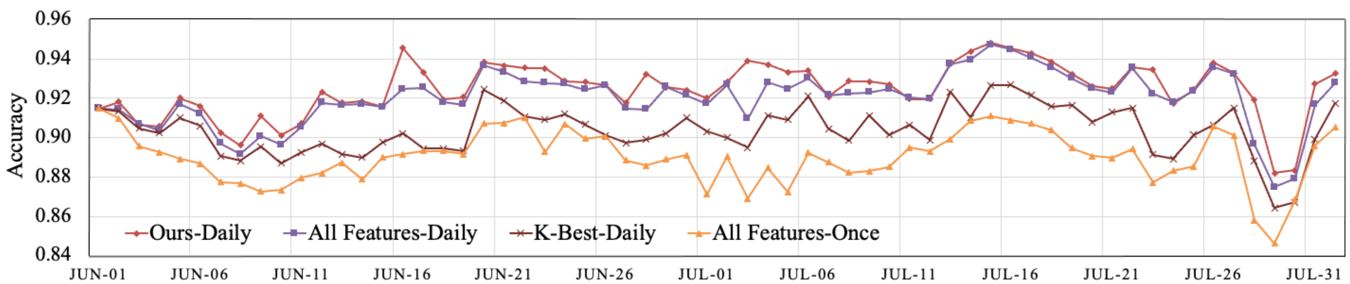

Figure 2: Daily classification accuracy using our method and three other baselines on the e-commerce fraud detection task.

terminates as the agent has skipped the last feature.

Reward with Curiosity

Environment Reward In order to optimize the FS objective directly, we construct the environment reward as the evaluating model performance on selected data, i.e.,

$$r_t^e = \text{Score}(\boldsymbol{D}[I_{t+1}]), \tag{1}$$

where $\text{Score}(\cdot)$ denotes the FS evaluation criterion and $\boldsymbol{D}[I_{t+1}]$ denotes the subset of dataset corresponding to I_{t+1}.

Noting that Eq. (1) does not rely on any concrete form of $\text{Score}(\cdot)$, our proposed framework could accomplish tailored selection according to the extensive demands of arbitrary FS performance metrics.

Intrinsic Reward by Curiosity The environment reward signal is given only at the end of each episode, before which there could be a long trajectory of zero-rewarded steps. Consequently, we adopt the self-supervised intrinsic curiosity module to encourage the exploration of novel states (Pathak et al. 2017). The error between the current state encoding $\phi(s_t)$ and the predicted state encoding $\hat{\phi}(s_t)$, scaled by factor $\eta = 0.01$, serves as our extra intrinsic reward, i.e.,

$$r_t^i = \frac{\eta}{2} \left\| \hat{\phi}(s_{t+1}) - \phi(s_{t+1}) \right\|_2^2. \tag{2}$$

Experiment

We compare our method with three other baselines facing feature drifts on an e-commerce fraud detection dataset. The classification dataset is collected in stream on a real-world e-commerce platform, containing around 20 million samples with 110 features, and each of its samples is a transaction labeled with 0 or 1, denoting non-fraud and fraud respectively.

The results are shown in Fig. 2. We could conclude from the results that our method (the red line) tracks effective features every day and outperforms the conventional feature selection methods, i.e., the K-best. And when the concept drift is quite intense, i.e., around JUL-29, our proposed method alleviates the performance decrease by preparing timely effective features for the downstream model.

Conclusion

In this paper, we have proposed an RL-based FS method dynamically adapting its selection strategy to feature drifting streams. Specifically, we design a skip-mode RL environment and further bonus the exploration with the ICM. With our work successfully deployed in the real-world business scenario, we expect it enlightens future RL solutions to a broader class of combination optimization problems.

References

Liu, K.; Fu, Y.; Wang, P.; Wu, L.; Bo, R.; and Li, X. 2019. Automating feature subspace exploration via multi-agent reinforcement learning. In *SIGKDD*, 207–215.

Pathak, D.; Agrawal, P.; Efros, A. A.; and Darrell, T. 2017. Curiosity-driven exploration by self-supervised prediction. In *ICML*, 2778–2787.