

Jiayu Li, Ziyi Ye, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma
 Information Retrieval Group, Department of Computer Science & Technology
 Tsinghua University

❖ Task Definition

➤ Recognize 20 types of Micro-activities

Ranking Task

- **Input:** query Q_i – Finding activity i
- **Model:** Similarity-based ranking model
- **Results:** Ranking list of instances I_j according to similarity $P(I_j | Q_i)$

Classification Task

- **Input:** Instance I_j – Classify instance to a type of activity
- **Model:** Multi-level classifier
- **Results:** Probability that I_j belongs to activity i Q_i $P(Q_i | I_j)$

➤ Task Equivalency:

$$P(I_j | Q_i) = \frac{P(Q_i | I_j)P(I_j)}{\sum_{k=1}^n P(Q_i | I_k)P(I_k)} \propto P(Q_i | I_j)$$

❖ 1. Feature Extraction

➤ Visual Features:

- Photos histogram similarity between instance & in instance
- Concepts and tags from multiple object detection methods

➤ OCR Features:

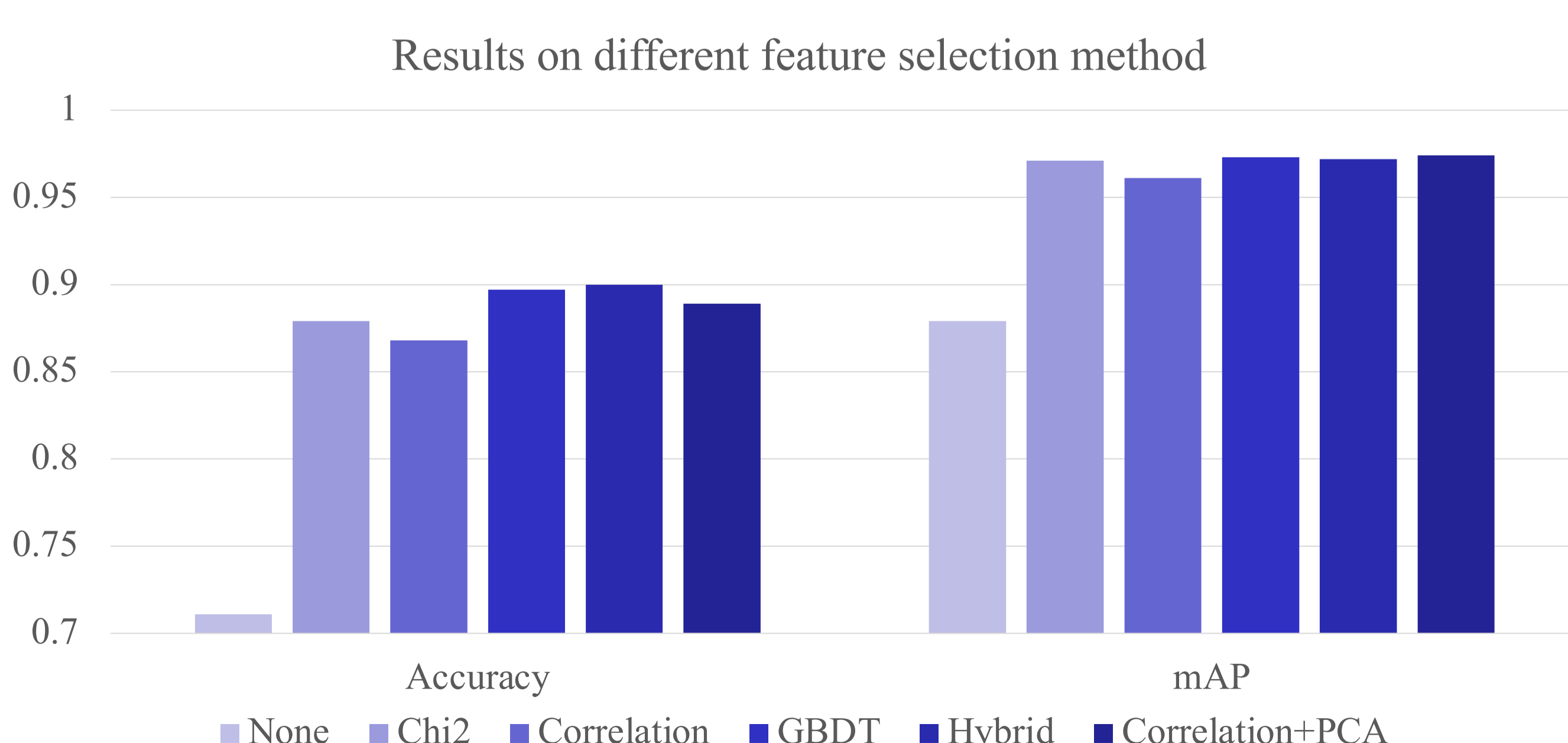
- URL extraction from screenshot
- OCR detection of the URL and head

➤ Temporal Features:

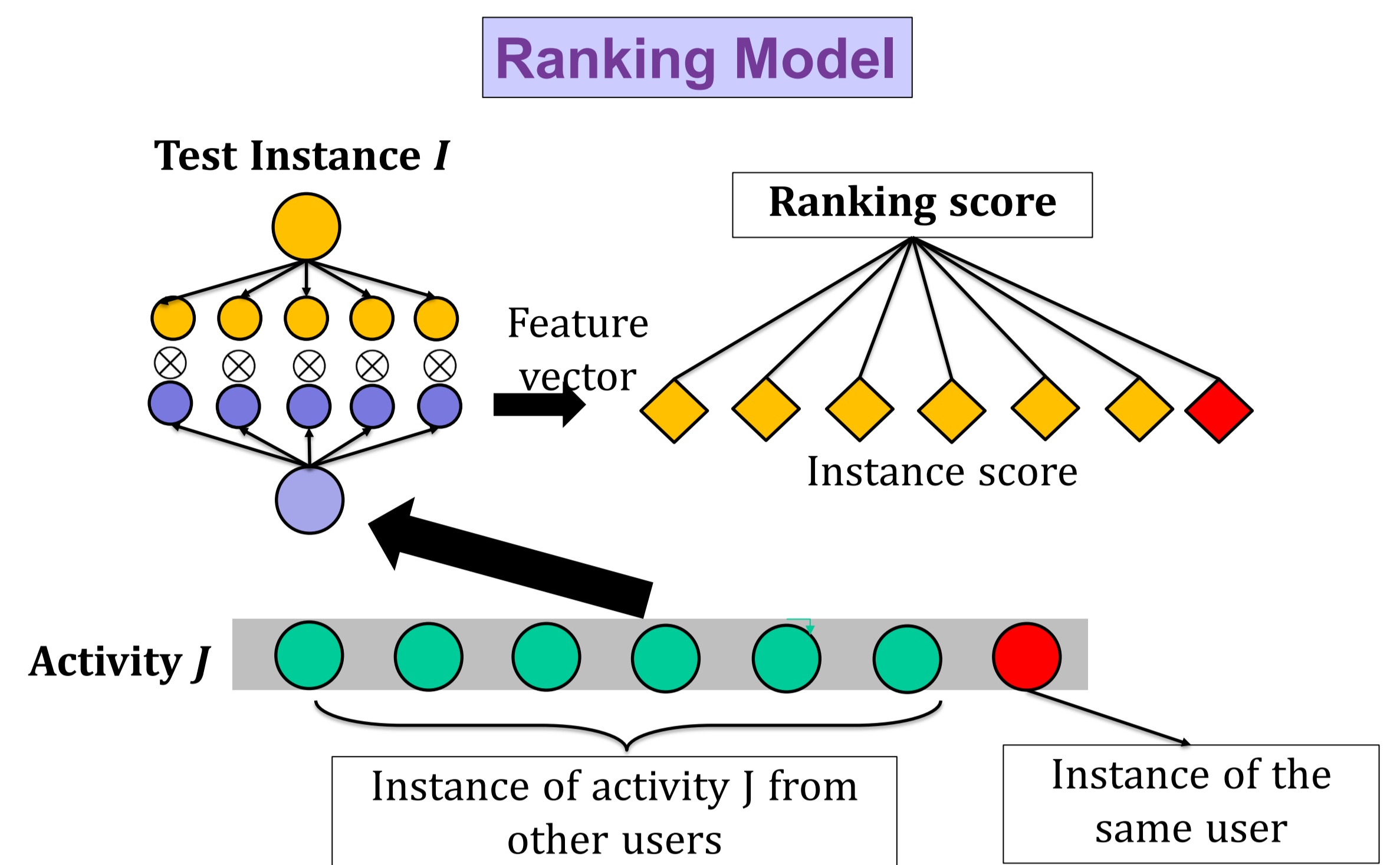
Feature type	Extraction Approaches
EOG	Time-domain statistic + Frequency-domain energy +LPC coefficients -> 62 features
Acceleration	3 position * 3 axes * Time-domain statistic + frequency-domain spectrum + axis correlation -> 523 features
Heart rate	Time-domain statistic -> 26 features
Mouse movement	Time-domain statistic + number of peaks -> 14 features

❖ 2. Feature Selection

- **Chi2:** Choose features with larger variance.
- **Correlation:** Remove similar features (high correlation)
- **GBDT:** Select features according to importance in GBDT
- **PCA:** Use Principal Component Analysis(PCA) to reduce dimension of visual features.
- **Hybrid:** Combine the above rules



❖ 3. Models



➤ Instance Score

- Similarity between Instance I and I' : $S_{I,I'} = \sum_{j=1}^5 \alpha_j * i_j \otimes i'_j$

➤ Ranking Score:

- Similarity between Instance I and Activity J :

$$S_{I,A} = \sum_{I' \in A} S_{I,I'} + \beta * \sum_{I'' \in A_u} S_{I,I''}$$

Classification Model

➤ Basic model selection:

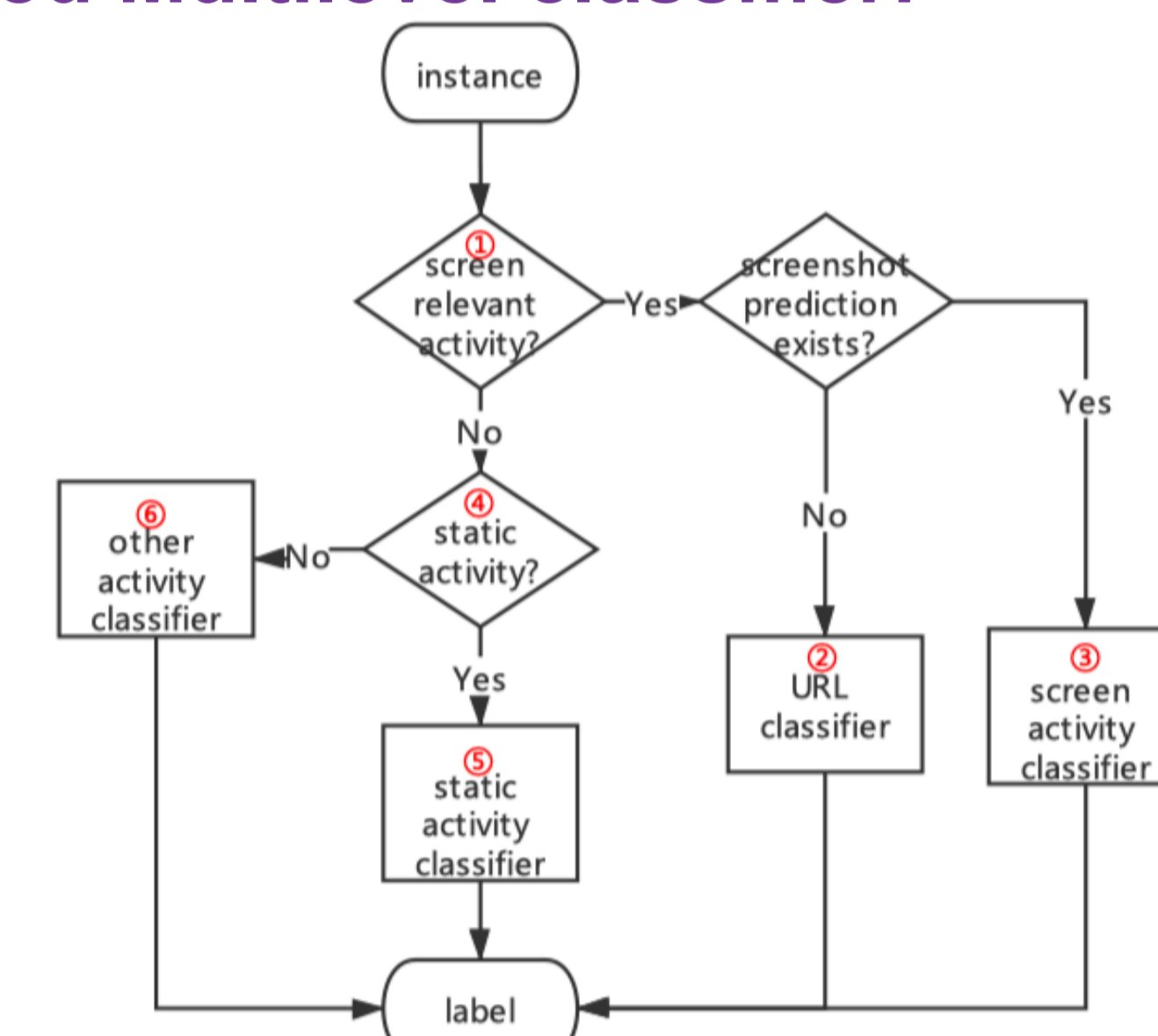
Classifier	LR	SVM	MLP	RF	XGboost	GBDT
Accuracy	0.825	0.821	0.811	0.779	0.826	0.836

- GBDT performs the best in all basic models.

➤ Rules detection:

- Partitioning the activities to find rules.
- Screen-relevant activities are similar, and can be classified with URL
- Static activities, i.e. Zoning out and closing eyes, are similar.

➤ Rule-based Multilevel classifier:



❖ 4. Results

Classifier	Accuracy	mAP (classify)	mAP (ranking)	Submission results
Basic GBDT Classifier	0.836	0.901	0.947	0.895
Ranking Model	0.789	0.843	0.836	0.782
Two-level Classifier (Cluster partition)	0.796	0.88	0.931	0.886
Two-level Classifier (Impurity partition)	0.875	0.921	0.971	0.901
Two-level Classifier (Similarity partition)	0.875	0.926	0.97	0.928
Rule-based Classifier	0.889	0.933	0.974	0.95

- Rule-based multi-level classifier performs the best