

都立高校生Twitterアカウントの 所属校推定の試み

全国高等学校情報教育研究会

全国大会(和歌山大会)

2019.08.11

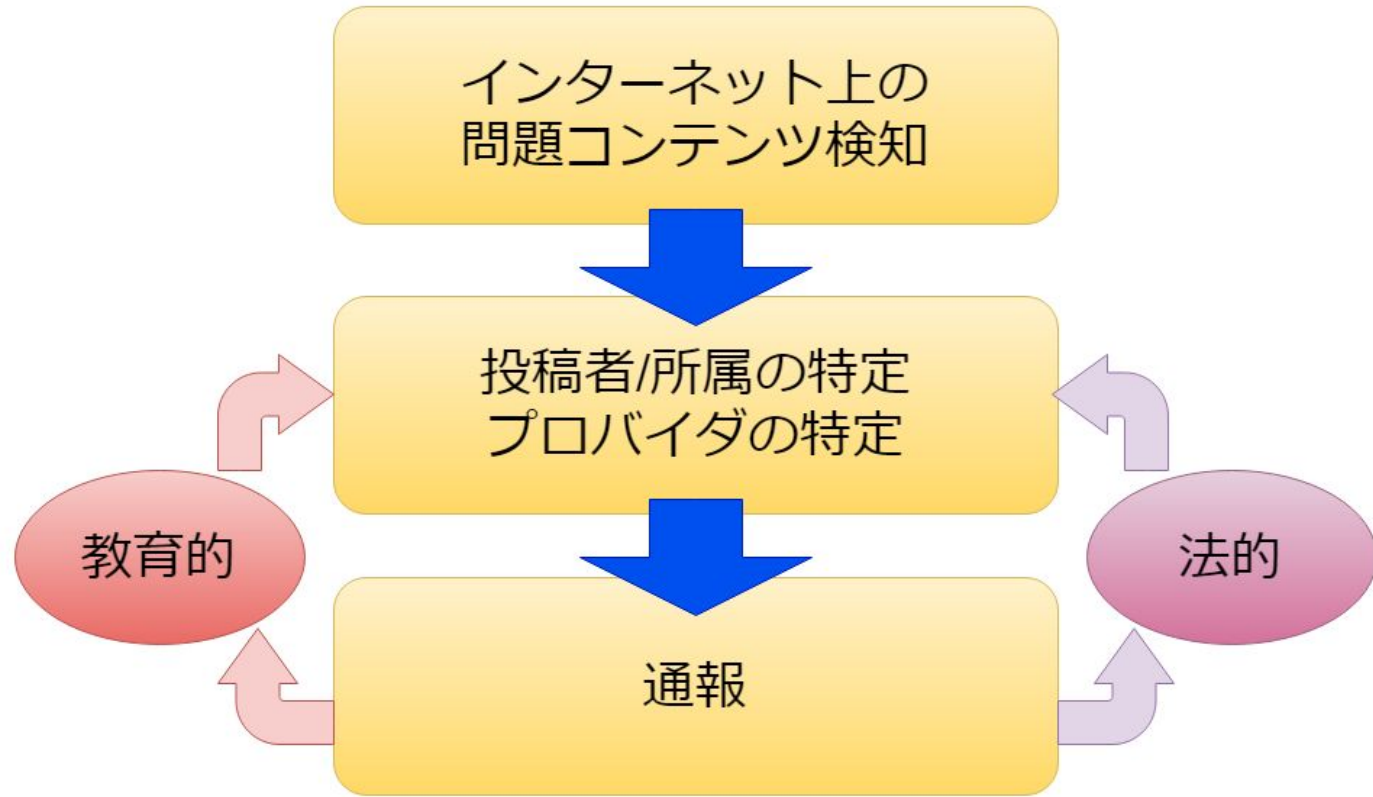
放送大学大学院修士課程 三村 徹(東京都教育委員会 指導主事)

放送大学 辰己丈夫

アジェンダ

- 1 研究の背景
- 2 サイバーパトロールの現状と課題
- 3 投稿者属性推定の技術動向
- 4 考案するアルゴリズム
- 5 今後の取組

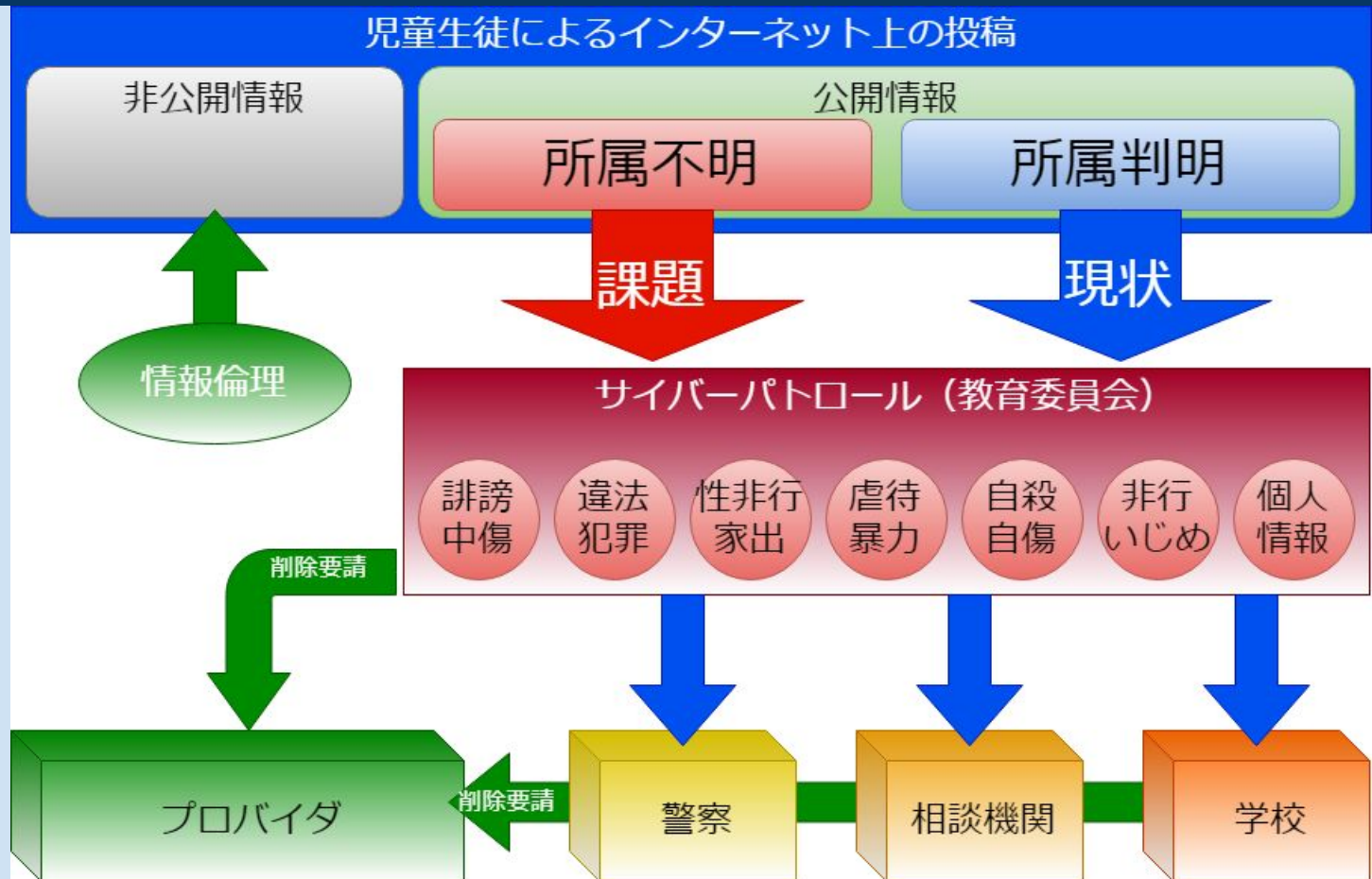
1 サイバーパトロールのシステムと機能



1 サイバーパトロールの必要性とネット利用の段階



2 サイバーパトロールの課題



2 サイバーパトロールの実施状況

【実施状況】

都道府県・指定都市 → 約70%が実施

市町村(指定都市除く) → 約10%が実施

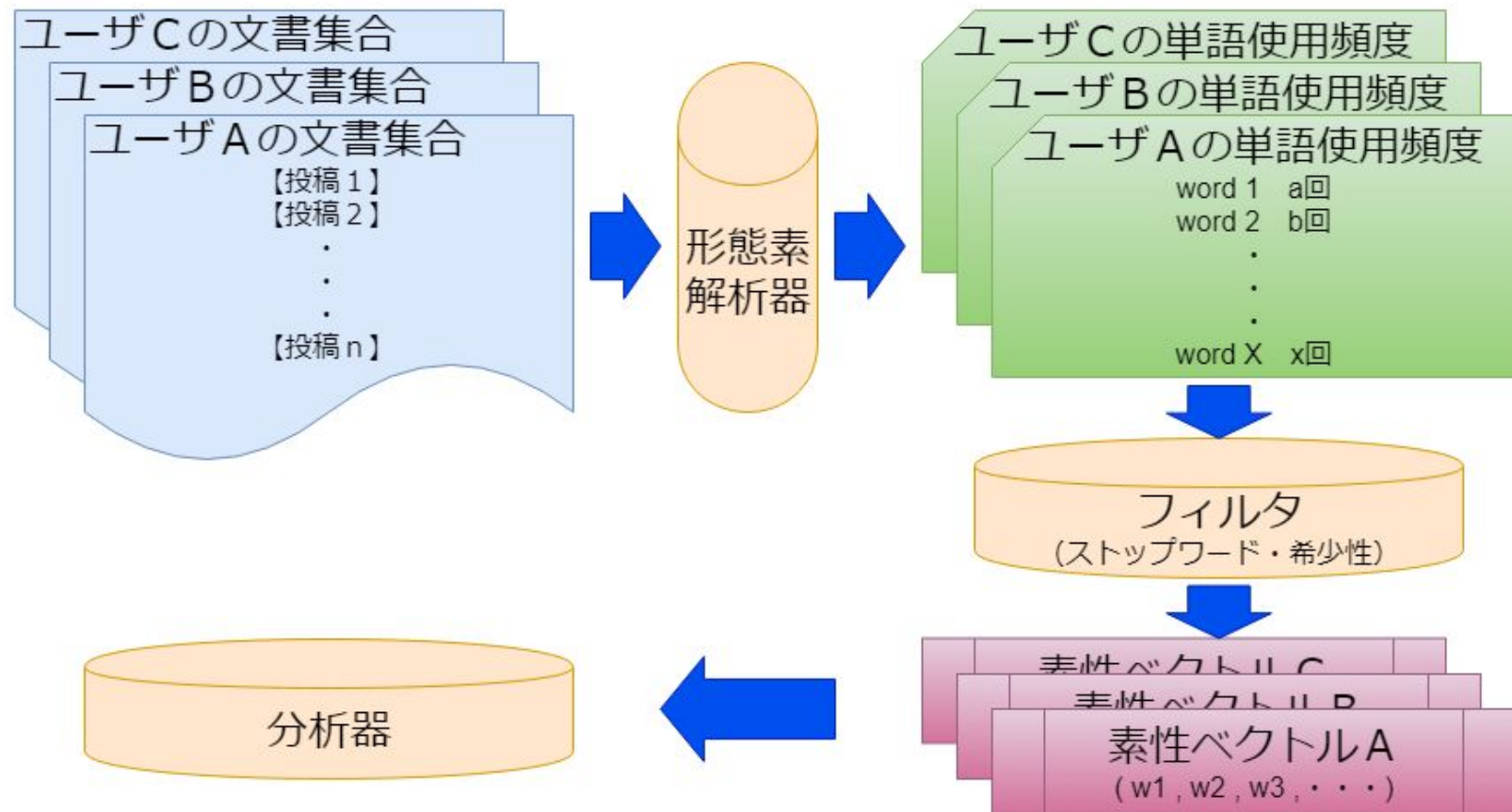
【実施体制】

都道府県・指定都市 → 約50%が外部委託で実施

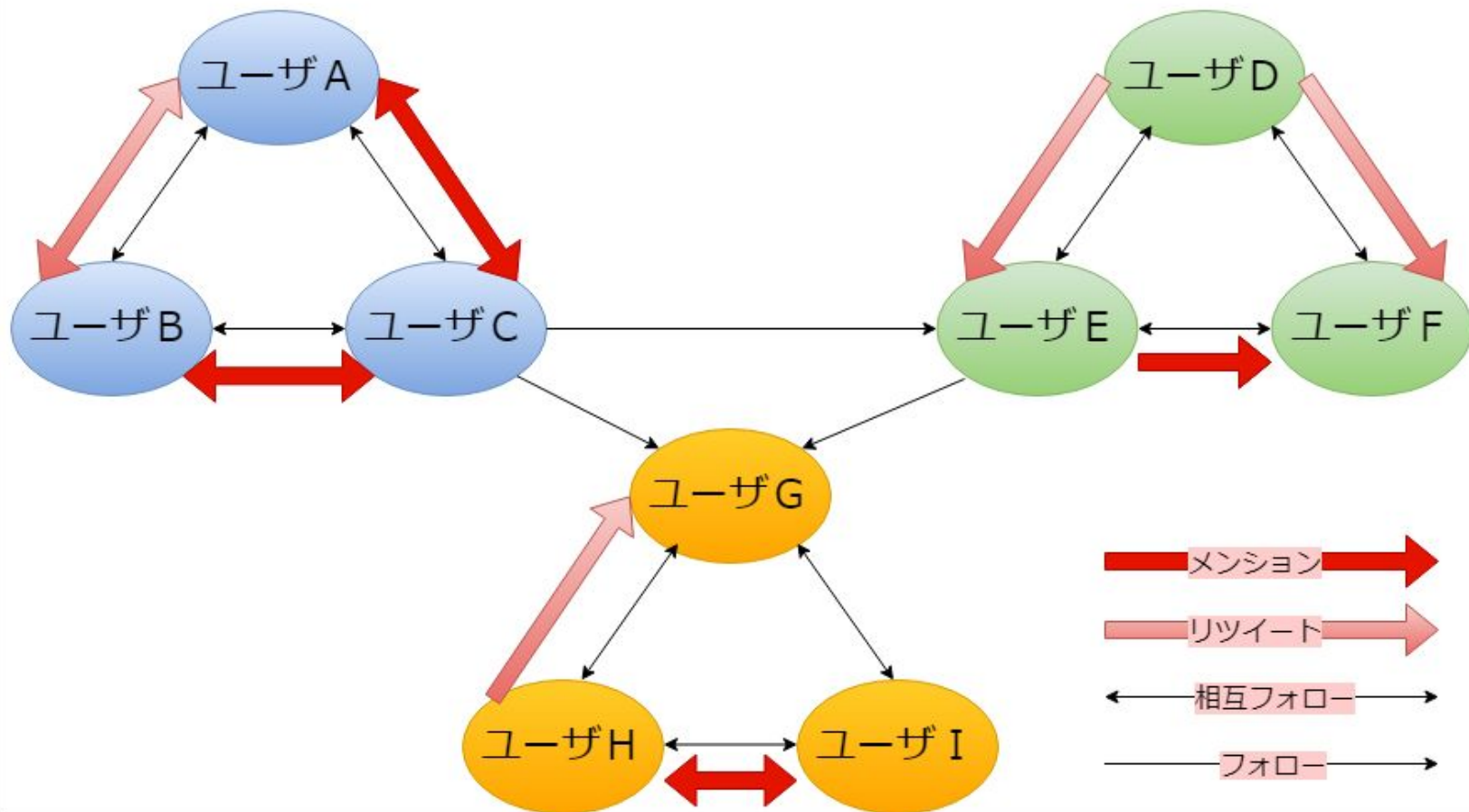
市町村(指定都市除く) → 約70%が内部職員で実施

費用・即時性・網羅性に課題

3 投稿者属性推定技術①「自然言語処理」



3 投稿者属性推定技術②「ソーシャルグラフ」



3 投稿者属性推定技術③「ジオタグ」

ジオタグ (geotag) : 緯度・経度の数値

例: $(n,e) = (35.689826, 139.691747)$

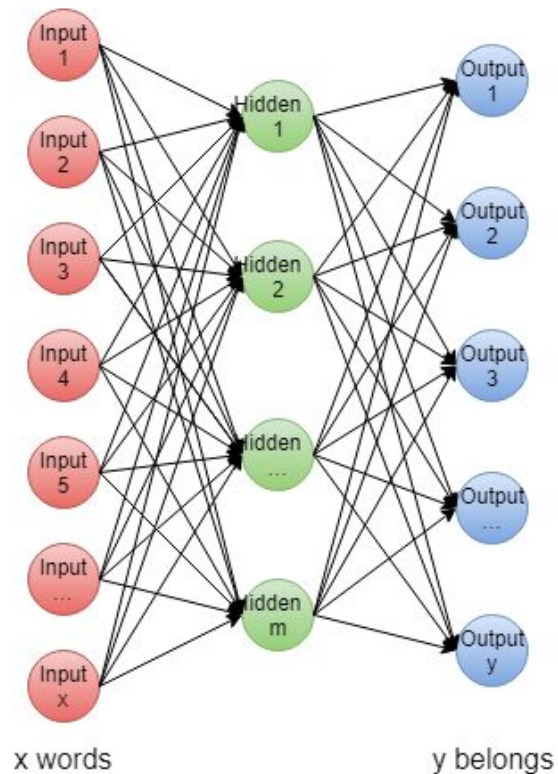
- ・投稿に付与するかは、利用者が選択
- ・地震や台風の自然災害の地域特定で有効
- ・曜日、時間帯と組み合わせ、職業推定で有効

【課題】

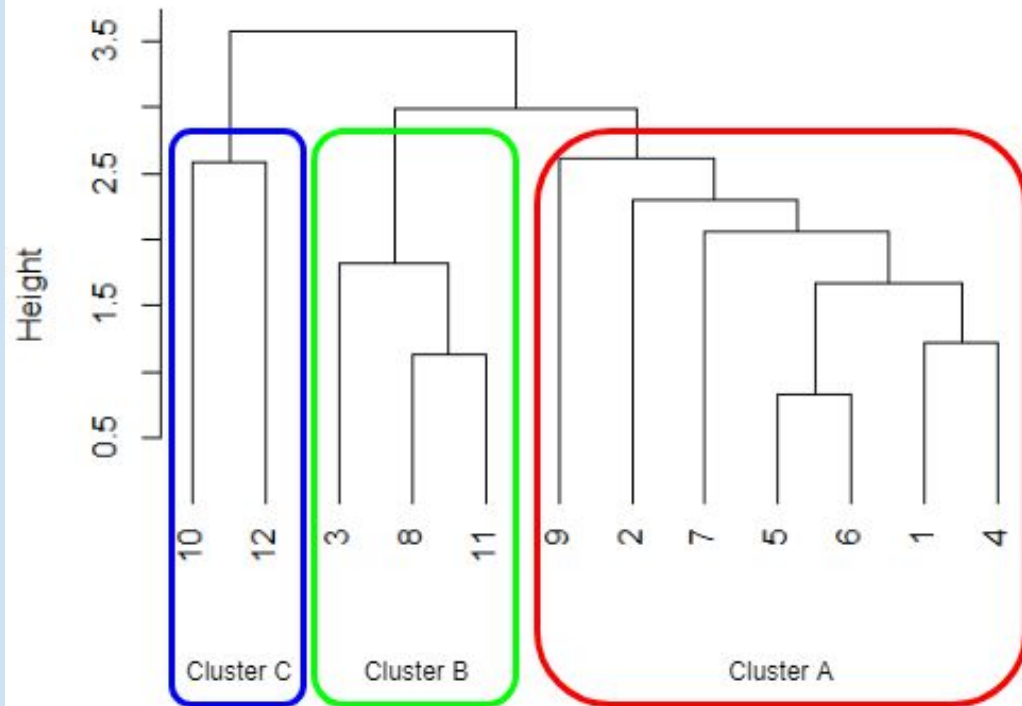
- ・ジオタグが付与された投稿が少数(約2%)

4 ニューラルネットワークと階層型クラスター分析

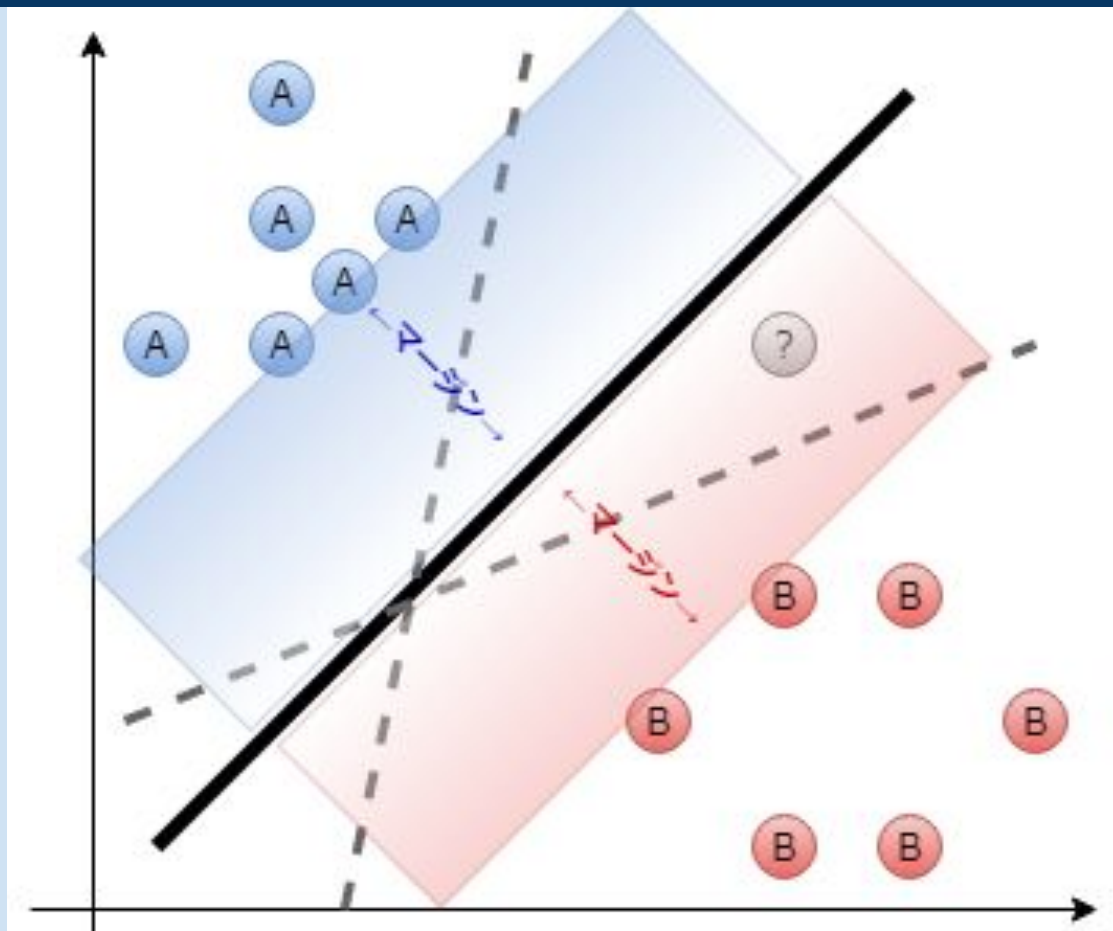
ニューラルネットワーク



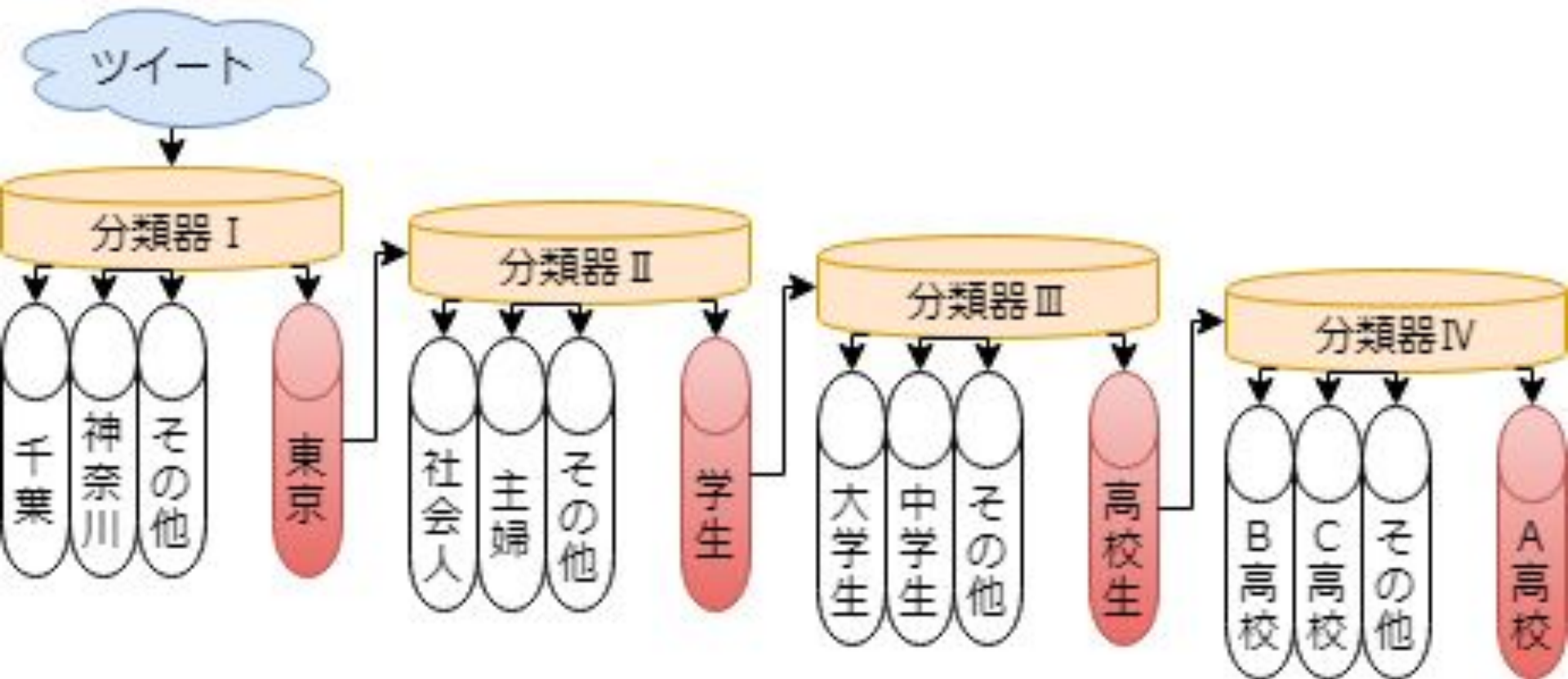
Cluster Dendrogram



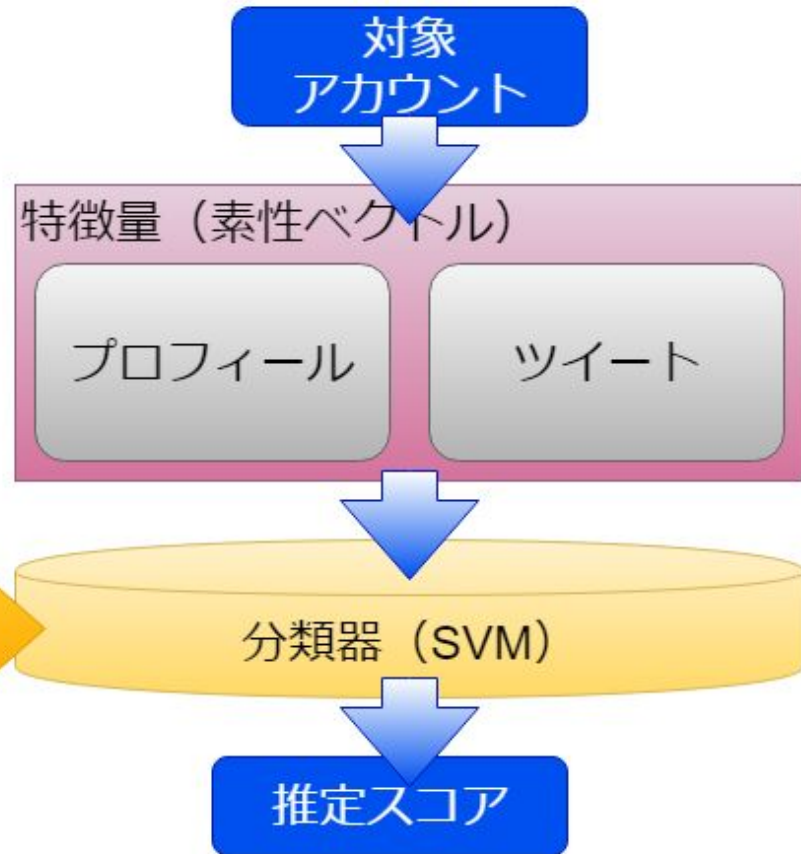
4 サポートベクトルマシン(SVM)



4 考案する手法「4層の分類器」



4 教師あり機械学習での分類方法(検討中)



5 今後の取組・展望

- ・東京都立高校を対象に試行
 - どの学校でも推定可能な技術として

- ・推定結果活用の課題整理
 - 推定結果(統計学的根拠)をどう扱うか？
 - プライバシーに配慮されているか？