

ユーザを意識した実データのマイニングと活用に関する研究

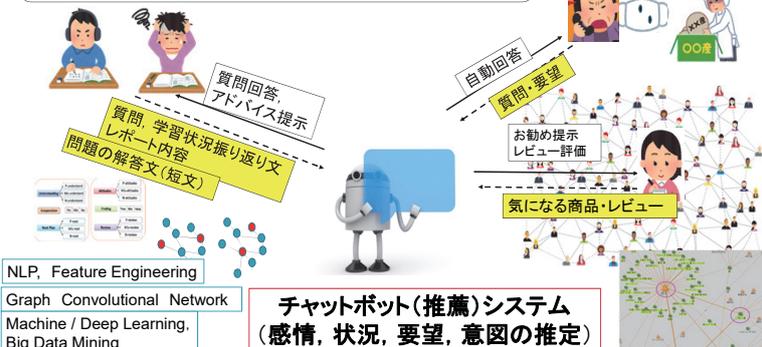
情報知能工学部門 高度ソフトウェア工学講座 准教授 峯 恒憲

▶ 様々な実データのマイニングと活用手法についての研究開発.

▶ 人間情報システム研究グループ (荒川・峯・福島研) <https://app.ait.kyushu-u.ac.jp/>
 峯研: <http://www.m.ait.kyushu-u.ac.jp/>



データマイニングと情報推薦

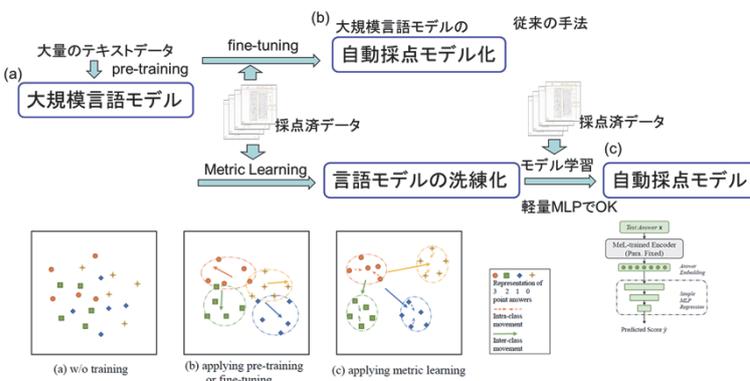


質問応答(チャットボット)システム

- 非質問文や領域外質問文などの識別
- 少ない教師データからの質問・応答対学習
- ユーザの感情推定・識別
 - 感情変化が頻繁におきる文脈を対象
- 有用レビュー文推定 (Amazonなど)
- Fake Newsの解析・自動推定
- コミュニティノートの各記事の信頼性推定
- 文脈と非定常性を考慮した情報推薦
- 多腕バンディッド, 実際の商用システムのデータ
- 固有表現抽出
 - 材料科学分野の特許文献等からの表現抽出
- 教育データマイニング
 - 成績推定, アドバイス文・報告文生成

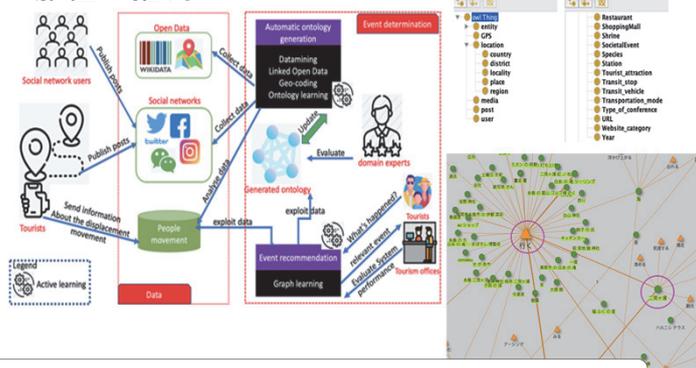


- 問題文の難易度推定
- 共通テストのデータを活用
- 短答記述式解答文の自動採点
 - 入試解答の自動採点に向けて(表現学習)
 - 学生の解答中の誤りに対する繰り返し指導



旅行情報の獲得と推薦

- 知識グラフの自動生成
- 位置(地名)情報に基づくreview文収集,
- 固有表現抽出, クラス名の出現確率, Wikidata等での自動補正, 利用者との対話による自動補正
- 旅行者には知られていない場所やイベントの発見
- 旅程の推薦



高度交通システム向けデータマイニング

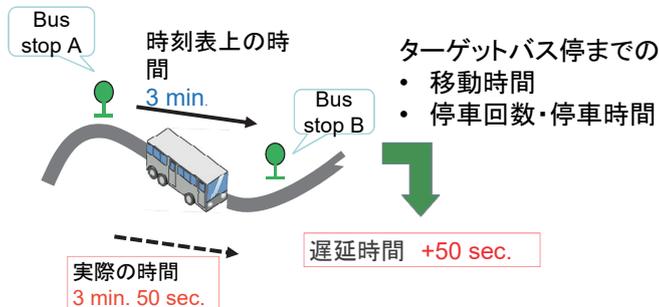
(1) 運行状況に関わる多様なデータ特徴の持続的収集・分析・調査



(2) 多様な目的達成(分析・推薦)に向けた高精度深層学習モデル開発と、その持続可能性担保のための機械学習基盤(MLOps)の構築

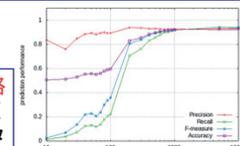
運行最適化 移動時間, 遅延時間, 停車時間, 停車回数	利用者満足度向上 乗り継ぎ可否, 遅刻回避, 満空情報などによる快適性確保	安全性向上 危険運転挙動, 事故発生可能性, 危険状況など
--	---	---

移動時間・到着時間・遅れ時間推定



事故抑止のための危険(急ブレーキ)状況推定
場所以外の情報(時間, 曜日, 方向, 速度など)の影響大

気象情報なしでも高確率で識別



Atmospheric pressure (気圧)	MO
Temperature (気温)	MO and AMeDAS
Humidity (湿度)	MO
Rainfall amount (雨量)	AMeDAS
Weather (天気)	MO
Wind speed (風速)	MO and AMeDAS
Wind direction (風向)	MO and AMeDAS
Visibility (視野)	MO

K	Precision	Recall	F1	Volume
ALL	0.9212	0.9431	0.9320	263M
10K	0.9211	0.9429	0.9319	263M

	MO	AMeDAS
# of positions (設置数)	19	57
Measurement Frequency (観測頻度)	60 min.	10 min.