



Санкт-Петербургский государственный университет
Кафедра системного программирования

Оценка личностных особенностей пользователя по данным страницы в Instagram

Александр Андреевич Федькин, группа 20.Б11-мм

Научный руководитель: к.т.н., доцент кафедры информатики СПбГУ Абрамов М.В.

Консультант: м.н.с. СПб ФИЦ РАН Корепанова А.А.

Санкт-Петербург
2022

- Instagram — одна из самых популярных социальных сетей
- Личность человека характеризуется пятью параметрами:
 - ▶ Экстраверсия
 - ▶ Привязанность
 - ▶ Самоконтроль
 - ▶ Эмоциональная устойчивость
 - ▶ Экспрессивность
- Полезно уметь предсказывать характер человека по его профилю:
знание характера \implies знание потребностей

Целью данной учебной практики является разработка модели оценки личностных особенностей пользователя по его профилю Instagram

Задачи:

- Произвести обзор существующих решений
- Разработать приложение с графическим интерфейсом для сбора данных из Instagram
- Собрать датасет: имя пользователя Instagram — результат психологического теста
- Обучить три модели, основанные на визуальных, социальных признаках и их комбинации
- Произвести сравнение моделей

Приложение для сбора данных

Критерии сравнения:

- Наличие нужной функциональности: выгрузка подписчиков, подписок, контента
- Поддержка создания цепочек запросов
- Наличие графического интерфейса
- Инструмент является бесплатным

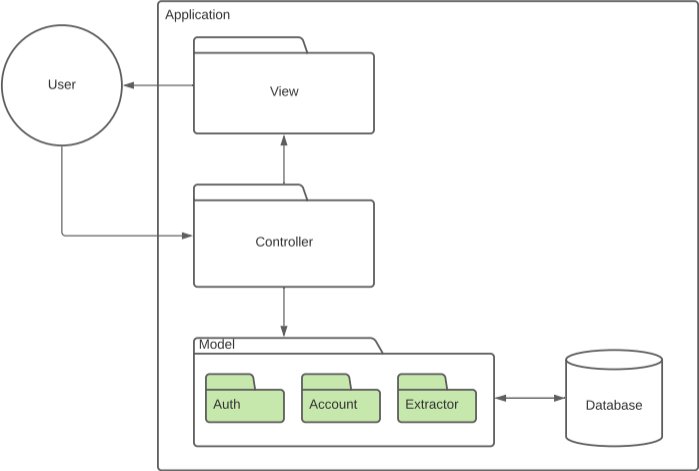
Параметр	phantombuster	instaloder	botster	apify
Функциональность	+	+	+	+
Цепочки запросов	-	-	-	-
Графический интерфейс	+	-	+	+
Бесплатный	-	+	-	-

Таблица: Сравнение инструментов для сбора данных из Instagram

Используемые инструменты:

- Python
- Instagrapi
- PyQt5
- Firebase Database
- PyTest
- PyLint

Архитектура приложения для сбора данных



Работы, посвященные предсказыванию черт характера пользователя по его профилю Instagram:

- Predicting Users' Personality from Instagram Pictures: Using Visual and/or Content Features? [1]
- Using computer vision techniques on Instagram to link users' personalities and genders to the features of their photos: An exploratory study [3]
- Fusing Social Media Cues: Personality Prediction from Twitter and Instagram [2]

Используемые инструменты:

- Python
- NumPy
- pandas
- OpenCV
- DeepFace
- scikit-learn

Нанято 279 респондентов для прохождения психологического теста

Критерии фильтрации:

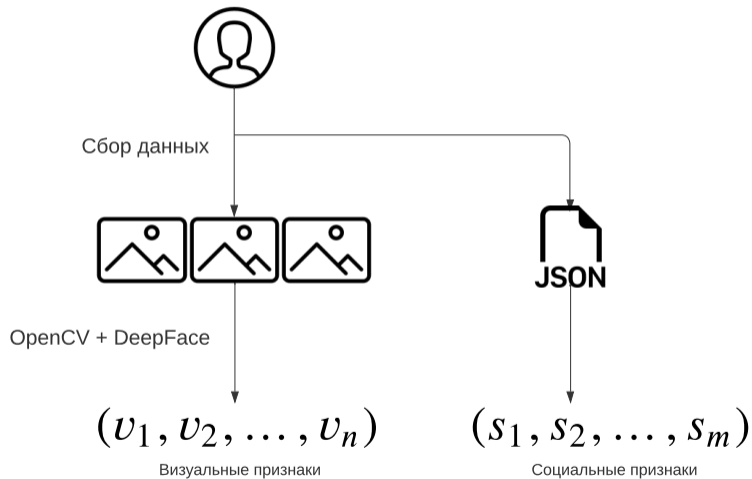
- Малое количество фотографий в профиле
- Большое количество фотографий рекламного характера/развлекательных изображений
- Нечестное прохождение опроса

Визуальные:

- Оттенок
- Насыщенность
- Яркость
- Количество лиц на фотографиях
- Эмоции на лицах:
 - ▶ Страх
 - ▶ Отвращение
 - ▶ Счастье
 - ▶ Грусть
 - ▶ Удивление
 - ▶ Нейтральность

Социальные:

- Число подписчиков
- Число подписок
- Число публикаций
- Число лайков
- Число комментариев






- Алгоритмы обучения: случайный лес, градиентный бустинг
- Подбор гиперпараметров: поиск по сетке
- Метрика: $RMSE = \sqrt{\frac{\sum_{i=1}^N (a(x_i) - y_i)^2}{N}} \in [1, \sqrt{24}]$
- Оценка качества: кросс-валидация с 10 итерациями

Черта характера	Алгоритм	Виз.	Соц.	Комб.
Экстраверсия	RF	0.7436	0.7511	0.7470
	GB	0.7702	0.8143	0.7542
Привязанность	RF	0.7054	0.7205	0.7093
	GB	0.7565	0.8163	0.7336
Самоконтроль	RF	0.6920	0.7155	0.6939
	GB	0.7388	0.7823	0.7067
Устойчивость	RF	0.8775	0.8684	0.8742
	GB	0.8957	0.9323	0.8765
Экспрессивность	RF	0.6767	0.6686	0.6671
	GB	0.6796	0.7263	0.6952

Таблица: Оценка качества моделей

- Произведен обзор существующих решений и аналогов
- Разработано приложение для сбора данных из Instagram
- Собран и отфильтрован датасет
- Обучены модели на социальных, визуальных и комбинированных признаках
- Произведено сравнение моделей

Реализация приложения для сбора данных доступна в [репозитории GitHub](#)

-  B. Ferwerda M. Tkalcic. Predicting Users' Personality from Instagram Pictures: Using Visual and/or Content Features? // PROCEEDINGS OF THE 26TH CONFERENCE ON USER MODELING, ADAPTATION AND PERSONALIZATION (UMAP'18). — 2018. — P. 157–161.
-  M. Skowron M. Tkallcic B. Ferwerda M. Schedl. Fusing Social Media Cues: Personality Prediction from Twitter and Instagram // PROCEEDINGS OF THE 25TH INTERNATIONAL CONFERENCE ON WORLD WIDE WEB (WWW'16 COMPANION). — 2016. — P. 107–108.
-  Y. Kim J. Kim. Using computer vision techniques on Instagram to link users' personalities and genders to the features of their photos: An exploratory study // INFORMATION PROCESSING & MANAGEMENT. — 2018. —

Дополнительный слайд: анализ датасета

