



مقایسه‌ی فالوئرهای دو فرد شاخص در توئیتر

مرحله ۱) جمع آوری دیتا

مرحله ۱ (جمع آوری دیتا



محمد جواد اکبرین



حمید رسایی

مرحله ۱ (جمع‌آوری دیتا)



محمد جواد اکبرین

آخرین توییت‌ها



حمید رسایی

آخرین توییت‌ها

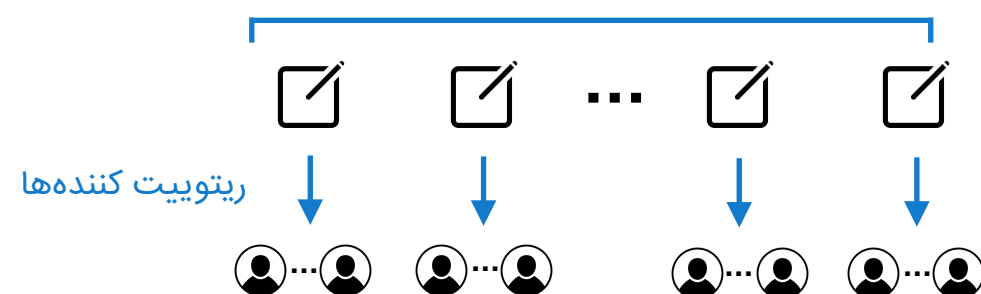


مرحله ۱ (جمع‌آوری دیتا)



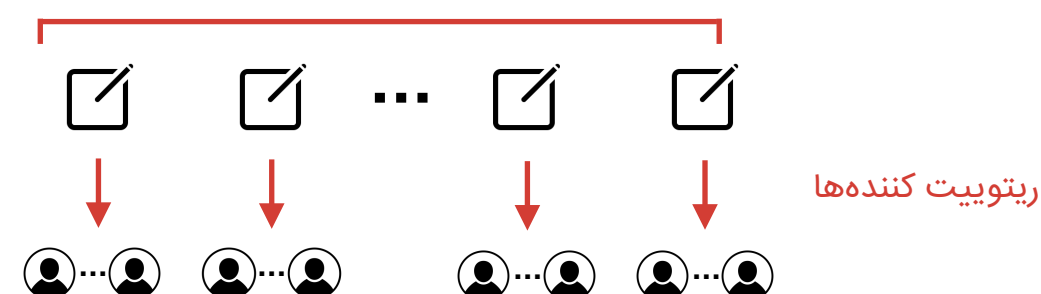
محمد جواد اکبرین

آخرین توییت‌ها



حمید رسایی

آخرین توییت‌ها

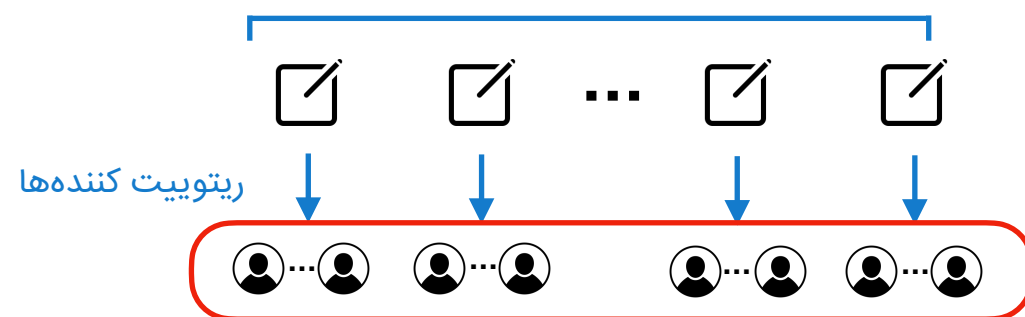


مرحله ۱ (جمع‌آوری دیتا)



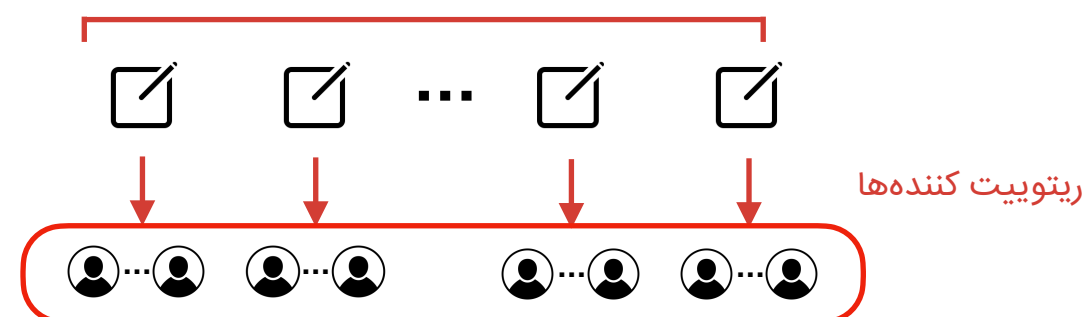
محمد جواد اکبرین

آخرین توییت‌ها



حمید رسایی

آخرین توییت‌ها

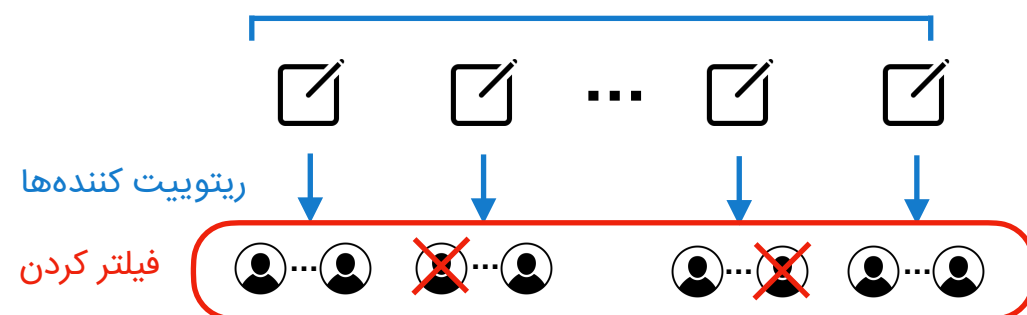


مرحله ۱) جمع‌آوری دیتا



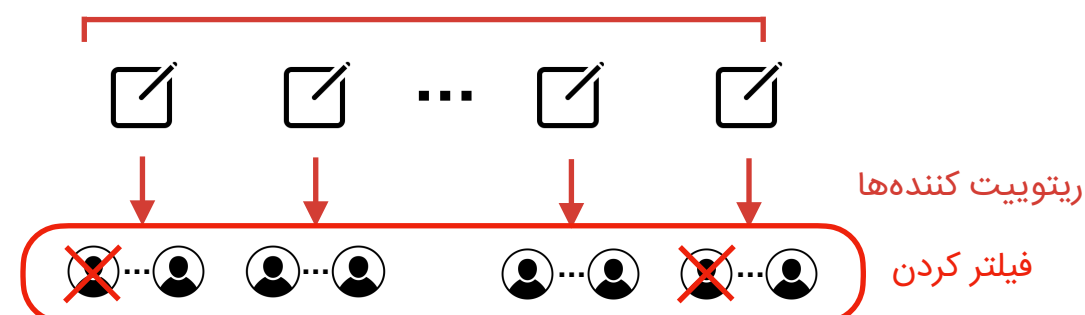
محمد جواد اکبرین

آخرین توییت‌ها



حمید رسایی

آخرین توییت‌ها

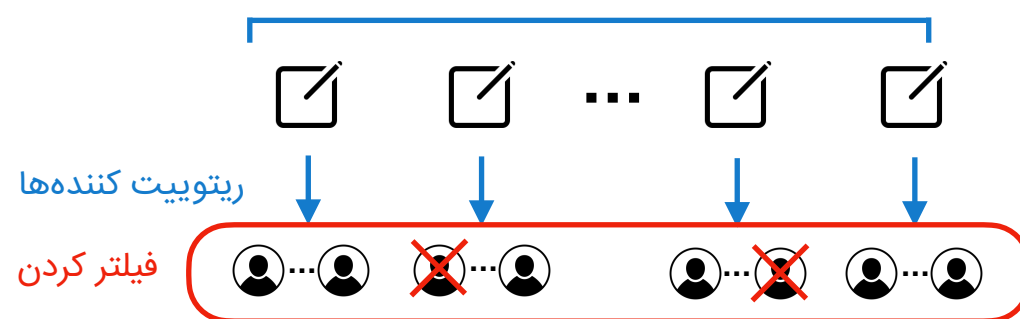


مرحله ۱) جمع‌آوری دیتا

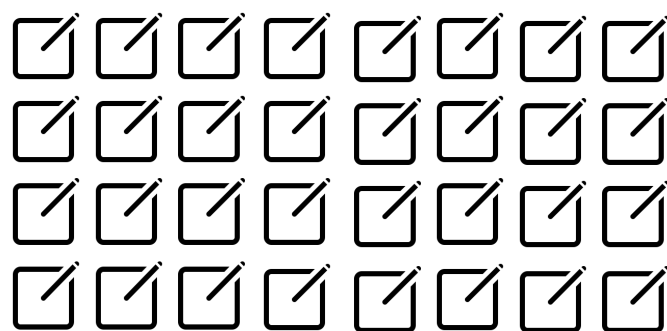


محمد جواد اکبرین

آخرین توییت‌ها

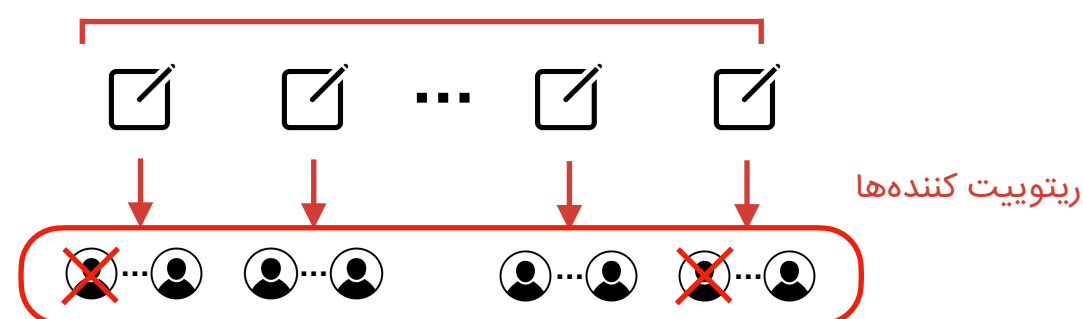


آخرین توییت‌های هر کاربر

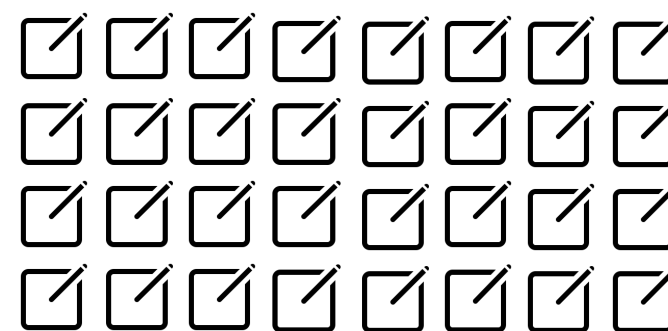


حمید رسایی

آخرین توییت‌ها



آخرین توییت‌های هر کاربر

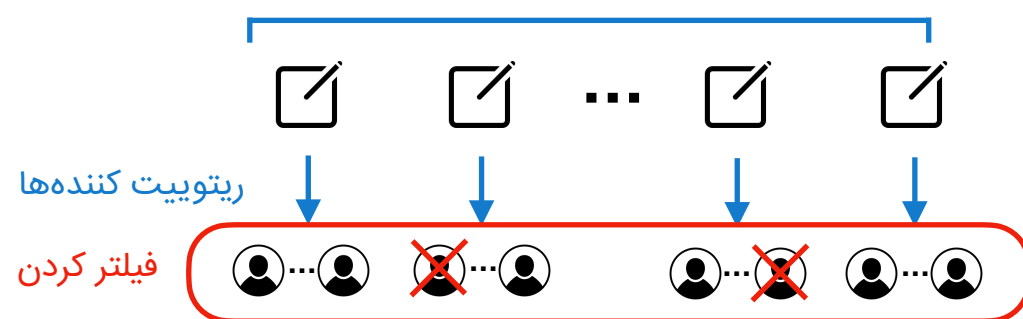


مرحله ۱) جمع‌آوری دیتا



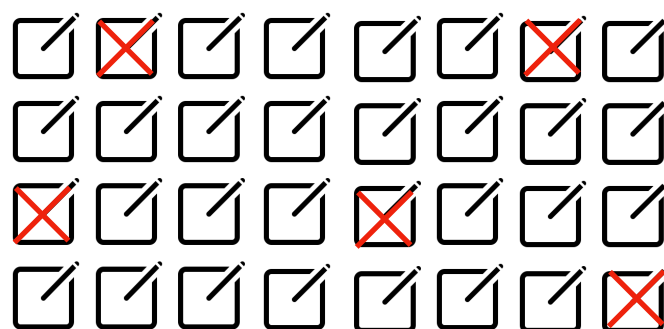
محمد جواد اکبرین

آخرین توییت‌ها



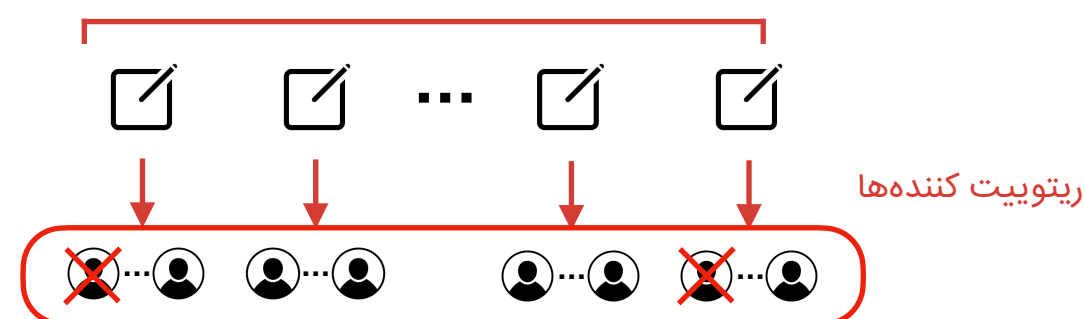
آخرین توییت‌های هر کاربر

فیلتر کردن



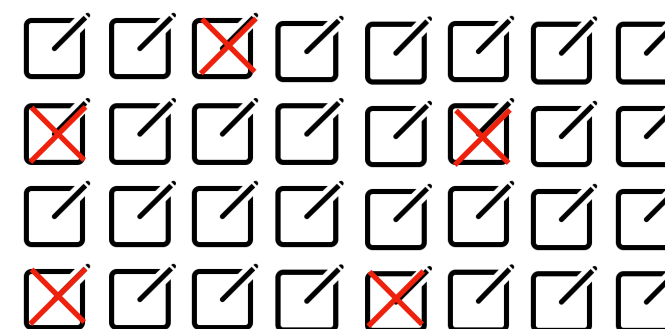
حمید رسایی

آخرین توییت‌ها



آخرین توییت‌های هر کاربر

فیلتر کردن



مرحله ۱ (جمع‌آوری دیتا)



محمد جواد اکبرین



توییت‌ها



حمید رسایی



توییت‌ها

مرحله ۱ (جمع‌آوری دیتا)



محمد جواد اکبرین



توییت‌ها



{token1, token2, ..., tokenN}



حمید رسایی



توییت‌ها



{token1, token2, ..., tokenN}

دیتا از لحاظ آماری

حمید رسایی

محمدجواد اکبرین

/

4,951,104	18,194,969	توکن
283,051	644,339	توکن یکتا
459,280	1,728,339	توییت

مرحله ۲ (تمیزکردن دیتا

مرحله ۲) تمیزکردن دیتا

- حذف لینک‌ها

مرحله ۲) تمیزکردن دیتا

- حذف لینک‌ها
- حذف کاراکترهای انگلیسی

مرحله ۲) تمیزکردن دیتا

- حذف لینک‌ها
- حذف کاراکترهای انگلیسی
- تبدیل کاراکترهای عربی به معادل فارسی

مرحله ۲) تمیزکردن دیتا

- حذف لینک‌ها
- حذف کاراکترهای انگلیسی
- تبدیل کاراکترهای عربی به معادل فارسی
- تبدیل ارقام عربی به فارسی

مرحله ۲) تمیزکردن دیتا

- حذف لینک‌ها
- حذف کاراکترهای انگلیسی
- تبدیل کاراکترهای عربی به معادل فارسی
- تبدیل ارقام عربی به فارسی
- چسباندن «ها» و «می»

مرحله ۲) تمیزکردن دیتا

- حذف لینک‌ها
- حذف کاراکترهای انگلیسی
- تبدیل کاراکترهای عربی به معادل فارسی
- تبدیل ارقام عربی به فارسی
- چسبانندن «ها» و «می»
- حذف همهی نیم‌فاصله‌ها

مرحله ۳ (نتایج



گروه محمدجواد اکبرین

- ایران
- روحانی
- کار
- خوب
- دست
- خدا
- زندگی
- دوست
- رای





کلمات با تکرار برابر در هر دو گروه

- روحانی
- قانون
- انتخابات
- کشور
- پول
- زن
- مشکل
- دل
- سال



کلمات با بیشترین اختلاف نسبی در گروه حمید رسایی

- برجام
- آمریکا
- امام
- شهید
- الله
- رهبری
- داعش
- ترامپ
- فتنه
- مجلس
- اسلام
- حمایت
- بنده
- حناب
- حضرت



کلمات با بیشترین اختلاف نسبی در گروه محمدجواد اکبرین

- حال
- خون
- حکومت
- جمهوری
- تویتر
- نگاه
- فیلم
- جامعه
- درد
- سیاسی
- دختر
- زندگی



با تشكر از توجه شما