



Привет, дорогой %Username% !

Для Мира настали темные времена. Поглощая одного пользователя за другим, социальные сети продолжают наращивать свою мощь. Но находятся малочисленные смельчаки, готовые укротить силу социальных сетей. Data Scientist-ами и Big Data Engineer-ами их все обычно называют. Их очень мало, но они есть. Мы – компания InData Labs – среди них.

Лучший способ вступить¹ в наши ряды – записаться в **R&D-лабораторию компании InData Labs и НИИ ППМИ**

БГУ. Но сперва тщательный Конкурсный отбор преодолеть Ты должен. Для этого:

1. Покажи себя – прими участие в Состязании!
2. Пришли нам **до 11:59:59 9 января** свое решение задачи на lab@indatalabs.com. Не стесняйся и не забудь рассказать также в письме о своих достижениях, навыках, умениях, интересах и личных целях в областях математики, программирования, Data Science и Computer Science.
3. Пройди устное собеседование.

Условие Состязания

Мир погрузился в Великую информационную войну. Одной из самых популярных сетей социальных Twitter является. Сообщение в ней потенциально опасным назовем, если оно верифицированным пользователем написано было и набрать за 48 часов более 20 ретвитов способно. Многие из таких сообщений огромную скрытую силу в себе таят.

Слабое место наш лазутчик в Twitter-е создать смог. Streaming API мы это место называем. Через него все сообщения получить мы можем, как только они были написаны. Модель раннего обнаружения потенциально опасных твитов разработать Тебе нужно. Data Science-силой наполнить ее следует.

Для построения, обучения и валидации модели фрагмент собранных нами исторических данных Тебе доступен. Количества ретвитов сообщений в нем соответствуют показателям через 48 часов после публикации. Скачать данные вместе с примерами входных и выходных файлов, необходимых для оценки Твоей работы, необходимо тут: <https://yadi.sk/d/3Mw6jo6Ej6cMx>. **Правила Состязания в файле README.pdf прочитать не забудь!**



Помня последствия атаки на Перл-Харбор, качество модели Твоей площадью под ROC-кривой измерять будут².

Присылай свою лучшую модель, ее исходный код и отчет о проведенном исследовании в формате PDF на lab@indatalabs.com.

Успехов! %Username%, мы верим в Тебя! Да пребудет с тобой Data Science!

¹ <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/> :-)

² <http://www.machinelearning.ru/wiki/index.php?title=ROC-кривая>
<https://ru.wikipedia.org/wiki/ROC-кривая>