






















# Программное средство для отслеживания и анализа показателей загрязнения окружающей среды на основе нейронной сети.


## Логи работы:

```
oap-spark-app_1 2019-06-01 16:25:50 INFO BlockManagerInfo:54 - Added broadcast_2277_piece0 in memory on fa9ab532a90b:38243 (size: 56.0 KB, free: 1048.7 MB)
oap-spark-app_1 2019-06-01 16:25:50 INFO SparkContext:54 - Created broadcast 2277 from broadcast at DAGScheduler.scala:1039
oap-spark-app_1 2019-06-01 16:25:50 INFO DAGScheduler:54 - Submitting 1 missing tasks from ResultStage 2277 (MapPartitionsRDD[11388] at save at SparkApp.scala:68)
first 15 tasks are for partitions Vector(0))
oap-spark-app_1 2019-06-01 16:25:50 INFO TaskSchedulerImpl:54 - Adding task set 2277.0 with 1 tasks
oap-spark-app_1 2019-06-01 16:25:50 INFO TaskSetManager:54 - Starting task 0.0 in stage 2277.0 (TID 2277, localhost, executor driver, partition 0, PROCESS_LOCAL, 39 bytes)
oap-spark-app_1 2019-06-01 16:25:50 INFO Executor:54 - Running task 0.0 in stage 2277.0 (TID 2277)
oap-spark-app_1 2019-06-01 16:25:50 INFO KafkaRDD:54 - Computing topic test_co-topic, partition 0 offsets 2338 -> 2340
oap-spark-app_1 2019-06-01 16:25:50 INFO CachedKafkaConsumer:54 - Initial fetch for spark-executor-consumer_id test_co-topic 0 2338
dumper 2019-06-01 16:25:50,240 [INFO] <air_pollution_dumper.CODumper> - Records was founded
dumper 2019-06-01 16:25:50,241 [INFO] <air_pollution_dumper.fs_adapter.DefaultFileSystem> - Makedir for path = ./output/co
dumper 2019-06-01 16:25:50,241 [INFO] <air_pollution_dumper.CODumper> - Dumping records
dumper 2019-06-01 16:25:50,241 [INFO] <air_pollution_dumper.fs_adapter.DefaultFileSystem> - Write to file = ./output/co/t4_132_y20182.json
dumper 2019-06-01 16:25:50,242 [INFO] <air_pollution_dumper.kafka_adapter.KafkaSender> - Send message to topic = test_co-topic; with key = /t4_132_y2018.3
n
oap-spark-app_1 2019-06-01 16:25:50 INFO FileOutputCommitter:108 - File Output Committer Algorithm version is 1
oap-spark-app_1 2019-06-01 16:25:50 INFO SQLHadoopMapReduceCommitProtocol:54 - Using output committer class org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
er
dumper 2019-06-01 16:25:50,251 [INFO] <air_pollution_dumper.kafka_adapter.KafkaSender> - RecordMetadata(topic='test_co-topic', partition=0, topic_partition=TopicPartition(topic='test_co-topic', partition=0), offset=2340, timestamp=1559406350242, checksum=988524794, serialized_key_size=18, serialized_value_size=21, serialized_header_size=-1)
dumper 2019-06-01 16:25:50,252 [INFO] <gcli> - Trying get data for topic = test_co-topic, latitude = 4, longitude = 32, year = 2018
dumper 2019-06-01 16:25:50,252 [INFO] <air_pollution_dumper.SODumper> - Trying find in file system
dumper 2019-06-01 16:25:50,252 [INFO] <air_pollution_dumper.SODumper> - Trying to fetch from server
oap-spark-app_1 2019-06-01 16:25:50 INFO FileOutputCommitter:535 - Saved output of task 'attempt_20190601162550_2277_m_000000_2277' to file:/shared/oap_out/_tempco
ry/0/task_20190601162550_2277_m_000000
oap-spark-app_1 2019-06-01 16:25:50 INFO SparkHadoopMapRedUtil:54 - attempt_20190601162550_2277_m_000000_2277: Committed
oap-spark-app_1 2019-06-01 16:25:50 INFO Executor:54 - Finished task 0.0 in stage 2277.0 (TID 2277). 2495 bytes result sent to driver
oap-spark-app_1 2019-06-01 16:25:50 INFO TaskSetManager:54 - Finished task 0.0 in stage 2277.0 (TID 2277) in 107 ms on localhost (executor driver) (1/1)
oap-spark-app_1 2019-06-01 16:25:50 INFO TaskSchedulerImpl:54 - Removed TaskSet 2277.0, whose tasks have all completed, from pool
oap-spark-app_1 2019-06-01 16:25:50 INFO DAGScheduler:54 - ResultStage 2277 (save at SparkApp.scala:68) finished in 0.133 s
oap-spark-app_1 2019-06-01 16:25:50 INFO DAGScheduler:54 - Job 2277 finished: save at SparkApp.scala:68, took 0.141810 s
oap-spark-app_1 2019-06-01 16:25:50 INFO FileFormatWriter:54 - Job null committed.
oap-spark-app_1 2019-06-01 16:25:50 INFO FileFormatWriter:54 - Finished processing stats for job null.
oap-spark-app_1 2019-06-01 16:25:50 INFO JobScheduler:54 - Finished job streaming job 1559406350000 ms.0 from job set of time 1559406350000 ms
oap-spark-app_1 2019-06-01 16:25:50 INFO JobScheduler:54 - Total delay: 0.292 s for time 1559406350000 ms (execution: 0.262 s)
oap-spark-app_1 2019-06-01 16:25:50 INFO KafkaRDD:54 - Removing RDD 11370 from persistence list
oap-spark-app_1 2019-06-01 16:25:50 INFO BlockManager:54 - Removing RDD 11370
oap-spark-app_1 2019-06-01 16:25:50 INFO ReceivedBlockTracker:54 - Deleting batches:
oap-spark-app_1 2019-06-01 16:25:50 INFO InputInfoTracker:54 - remove old batch metadata: 1559406340000 ms
dumper 2019-06-01 16:25:50,433 [INFO] <air_pollution_dumper.SODumper> - Records was founded
```

## Загруженные данные:

| /.../OpenAirPollutionMonitor/src/docker/zeppelin/data/oap_out/topic=test_co-topic/   |      |                    |
|--|------|--------------------|
| Name   | Size | Changed            |
|   |      | 6/1/2019 5:46:43 F |
|  part-00000-fab006e0-614e-4377-8cee-a417e8463...    | 1 KB | 6/1/2019 5:43:15 F |
|  part-00000-f50a70b5-cfb9-42ea-b4f1-66be98164...    | 1 KB | 6/1/2019 5:44:35 F |
|  part-00000-edc76bbe-14f1-4d34-bd7d-d130b567...     | 1 KB | 6/1/2019 5:44:26 F |
|  part-00000-ecba4c39-1e90-4986-b664-6d84b692...     | 1 KB | 6/1/2019 5:43:26 F |
|  part-00000-ea0c1b0c-79bc-4ae2-8e7f-c5e1e63e7...    | 1 KB | 6/1/2019 5:43:45 F |
|  part-00000-dd5682fc-b30e-43c7-bad4-2943c8705...    | 1 KB | 6/1/2019 5:43:06 F |
|  part-00000-d595076b-6f86-4e1e-afa0-39611b1ffd...   | 1 KB | 6/1/2019 5:44:30 F |
|  part-00000-ca894023-4678-4f8f-8eba-1ee1d8443...    | 1 KB | 6/1/2019 5:44:05 F |
|  part-00000-bd0c6e66-345b-40ec-a2cd-2e3b8a61...     | 1 KB | 6/1/2019 5:44:21 F |
|  part-00000-a8128b2b-ca5b-4ec4-8226-8b116722...     | 1 KB | 6/1/2019 5:43:30 F |
|  part-00000-a3e93250-3b42-4bb6-a55b-874e2619...     | 1 KB | 6/1/2019 5:44:00 F |
|  part-00000-7687780d-679c-4859-b0f1-542f4dc2d...    | 1 KB | 6/1/2019 5:43:21 F |
|  part-00000-51484f27-1712-404e-946d-df3a33371...    | 1 KB | 6/1/2019 5:43:51 F |
|  part-00000-15582f20-f547-453b-b1ec-9404d3c09...    | 1 KB | 6/1/2019 5:43:08 F |
|  part-00000-6384ea99-2ddd-4b1c-a7ed-451ca30f7...   | 1 KB | 6/1/2019 5:43:35 F |
|  part-00000-1250a88e-0e77-4fff-803f-9e34bd9c54... | 1 KB | 6/1/2019 5:43:03 F |
|  part-00000-934e952d-f7a8-41db-b400-05b37231e...  | 1 KB | 6/1/2019 5:44:15 F |
|  part-00000-924c88c7-06a7-4c9e-9f42-6c6780c7d...  | 1 KB | 6/1/2019 5:44:55 F |
|  part-00000-75dbe3af-c94f-4570-b5f4-9841aef616... | 1 KB | 6/1/2019 5:45:00 F |
|  part-00000-74df4db3-85f5-47fc-b12b-e2c3f25fd4... | 1 KB | 6/1/2019 5:43:40 F |

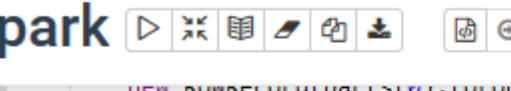
## Обработанные данные:

 Zeppelin


Notebook ▾Job

anonymous ▾

spark



Head ▾




```
}
new RowMatrix(pairs(0):colony, pairs(1):colony, pairs(2):colony, pairs(3):longitude)

// Read file as RDD
val topic = "test_co-topic"

val rdd=sc.textFile("/zeppelin/data/shared/oap_out/topic=" + topic + "/part-*.csv")
val rdd1 = rdd.map(parseRecord(_))

// Convert rdd to dataframe using toDF
val df = rdd1.toDF

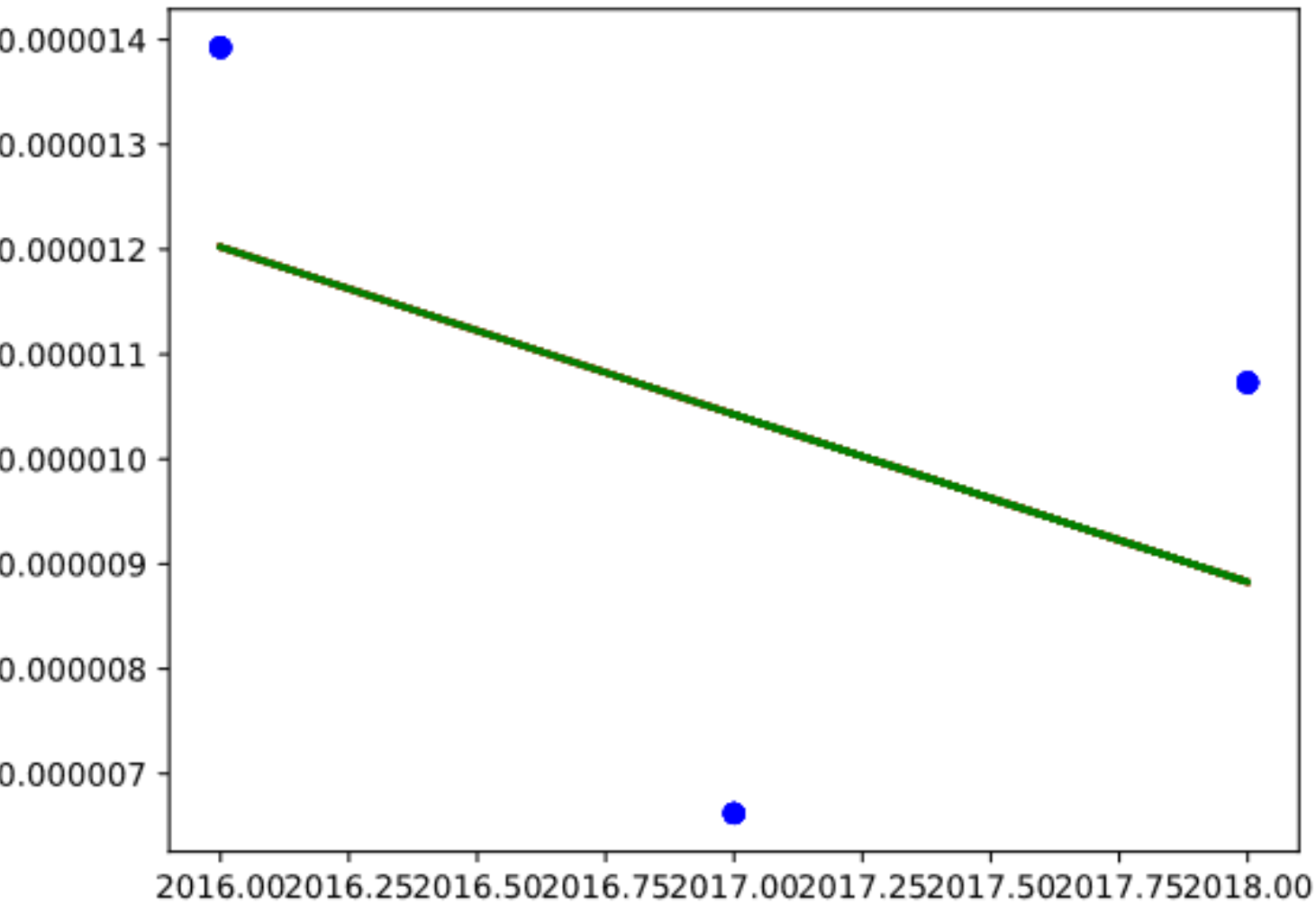
z.show(df)
```



settings ▾

| latitude | longitude | year | value                 |
|----------|-----------|------|-----------------------|
| 0        | 110       | 2018 | 1.1084645848313812E-5 |
| 0        | 111       | 2016 | 2.627907861096901E-6  |
| 0        | 56        | 2016 | 1.0720707905420568E-5 |
| 0        | 56        | 2017 | 1.3532550838135649E-5 |
| 0        | 143       | 2017 | 4.974422608938767E-6  |
| 0        | 143       | 2018 | 1.4041963368072174E-5 |
| 0        | 67        | 2018 | 1.5113662811927497E-5 |
| 0        | 68        | 2016 | 9.19997546589002E-6   |

## Результат регрессии:



|                                 |           |          |       |      |                 |       |          |
|---------------------------------|-----------|----------|-------|------|-----------------|-------|----------|
| ГУИР.400201.079 ПЛ.2            |           |          |       |      |                 |       |          |
| Заключительный плакат<br>Плакат |           |          |       |      | Лит.            | Масса | Масштаб  |
|                                 |           |          |       |      | У               |       |          |
|                                 |           |          |       |      | Лист            |       | Листов 1 |
|                                 |           |          |       |      |                 |       |          |
|                                 |           |          |       |      |                 |       |          |
|                                 |           |          |       |      |                 |       |          |
|                                 |           |          |       |      |                 |       |          |
| Изм.                            | Лист      | № докум. | Подп. | Дата | ЭВМ, фр. 550504 |       |          |
| Разраб.                         | Федоров   |          |       |      |                 |       |          |
| Прое.                           | Сасин     |          |       |      |                 |       |          |
| Т. контр.                       | Сасин     |          |       |      |                 |       |          |
| Реценз.                         |           |          |       |      |                 |       |          |
| Н. контр.                       | Сидорович |          |       |      |                 |       |          |
| Утв.                            | Никольшин |          |       |      |                 |       |          |