

Task 1

History Server - Google Chrome

Assign x My Cla x Workb x EMR - x Datap x EMR - x Histor x 10186 x 1.4.4 - x java - x LogMi x 15.7. lc x Python x +

Not secure | ec2-34-204-167-23.compute-1.amazonaws.com:18080

Spark 2.4.4 History Server

Event log directory: hdfs://varlog/spark/apps
Last updated: 2020-04-17 20:58:37
Client local time zone: America/New_York

Search:

App ID	App Name	Started	Completed	Duration	Spark User	Last Updated	Event Log
application_1587164871837_0001	Assignment5-task1	2020-04-17 19:09:58	2020-04-17 20:42:54	1.5 h	hadoop	2020-04-17 20:42:54	Download

Showing 1 to 1 of 1 entries
[Show incomplete applications](#)

Task 1

Task 2

Task 3

Task 4

Task 5

Task 6

Task 7

Task 8

Task 9

Task 10

Task 11

Task 12

Task 13

Task 14

Task 15

Task 16

Task 17

Task 18

Task 19

Task 20

Task 21

Task 22

Task 23

Task 24

Task 25

Task 26

Task 27

Task 28

Task 29

Task 30

Task 31

Task 32

Task 33

Task 34

Task 35

Task 36

Task 37

Task 38

Task 39

Task 40

Task 41

Task 42

Task 43

Task 44

Task 45

Task 46

Task 47

Task 48

Task 49

Task 50

Task 51

Task 52

Task 53

Task 54

Task 55

Task 56

Task 57

Task 58

Task 59

Task 60

Task 61

Task 62

Task 63

Task 64

Task 65

Task 66

Task 67

Task 68

Task 69

Task 70

Task 71

Task 72

Task 73

Task 74

Task 75

Task 76

Task 77

Task 78

Task 79

Task 80

Task 81

Task 82

Task 83

Task 84

Task 85

Task 86

Task 87

Task 88

Task 89

Task 90

Task 91

Task 92

Task 93

Task 94

Task 95

Task 96

Task 97

Task 98

Task 99

Task 100

Task 101

Task 102

Task 103

Task 104

Task 105

Task 106

Task 107

Task 108

Task 109

Task 110

Task 111

Task 112

Task 113

Task 114

Task 115

Task 116

Task 117

Task 118

Task 119

Task 120

Task 121

Task 122

Task 123

Task 124

Task 125

Task 126

Task 127

Task 128

Task 129

Task 130

Task 131

Task 132

Task 133

Task 134

Task 135

Task 136

Task 137

Task 138

Task 139

Task 140

Task 141

Task 142

Task 143

Task 144

Task 145

Task 146

Task 147

Task 148

Task 149

Task 150

Task 151

Task 152

Task 153

Task 154

Task 155

Task 156

Task 157

Task 158

Task 159

Task 160

Task 161

Task 162

Task 163

Task 164

Task 165

Task 166

Task 167

Task 168

Task 169

Task 170

Task 171

Task 172

Task 173

Task 174

Task 175

Task 176

Task 177

Task 178

Task 179

Task 180

Task 181

Task 182

Task 183

Task 184

Task 185

Task 186

Task 187

Task 188

Task 189

Task 190

Task 191

Task 192

Task 193

Task 194

Task 195

Task 196

Task 197

Task 198

Task 199

Task 200

Task 201

Task 202

Task 203

Task 204

Task 205

Task 206

Task 207

Task 208

Task 209

Task 210

Task 211

Task 212

Task 213

Task 214

Task 215

Task 216

Task 217

Task 218

Task 219

Task 220

Task 221

Task 222

Task 223

Task 224

Task 225

Task 226

Task 227

Task 228

Task 229

Task 230

Task 231

Task 232

Task 233

Task 234

Task 235

Task 236

Task 237

Task 238

Task 239

Task 240

Task 241

Task 242

Task 243

Task 244

Task 245

Task 246

Task 247

Task 248

Task 249

Task 250

Task 251

Task 252

Task 253

Task 254

Task 255

Task 256

Task 257

Task 258

Task 259

Task 260

Task 261

Task 262

Task 263

Task 264

Task 265

Task 266

Task 267

Task 268

Task 269

Task 270

Task 271

Task 272

Task 273

Task 274

Task 275

Task 276

Task 277

Task 278

Task 279

Task 280

Task 281

Task 282

Task 283

Task 284

Task 285

Task 286

Task 287

Task 288

Task 289

Task 290

Task 291

Task 292

Task 293

Task 294

Task 295

Task 296

Task 297

Task 298

Task 299

Task 300

Task 301

Task 302

Task 303

Task 304

Task 305

Task 306

Task 307

Task 308

Task 309

Task 310

Task 311

Task 312

Task 313

Task 314

Task 315

Task 316

Task 317

Task 318

Task 319

Task 320

Task 321

Task 322

Task 323

Task 324

Task 325

Task 326

Task 327

Task 328

Task 329

Task 330

Task 331

Task 332

Task 333

Task 334

Task 335

Task 336

Task 337

Task 338

Task 339

Task 340

Task 341

Task 342

Task 343

Task 344

Task 345

Task 346

Task 347

Task 348

Task 349

Task 350

Task 351

Task 352

Task 353

Task 354

Task 355

Task 356

Task 357

Task 358

Task 359

Task 360

Task 361

Task 362

Task 363

Task 364

Task 365

Task 366

Task 367

Task 368

Task 369

Task 370

Task 371

Task 372

Task 373

Task 374

Task 375

Task 376

Task 377

Task 378

Task 379

Task 380

Task 381

Task 382

Task 383

Task 384

Task 385

Task 386

Task 387

Task 388

Task 389

Task 390

Task 391

Task 392

Task 393

Task 394

Task 395

Task 396

Task 397

Task 398

Task 399

Task 400

Task 401

Task 402

Task 403

Task 404

Task 405

Task 406

Task 407

Task 408

Task 409

Task 410

Task 411

Task 412

Task 413

Task 414

Task 415

Task 416

Task 417

Task 418

Task 419

Task 420

Task 421

Task 422

Task 423

Task 424

Task 425

Task 426

Task 427

Task 428

Task 429

Task 430

Task 431

Task 432

Task 433

Task 434

Task 435

Task 436

Task 437

Task 438

Task 439

Task 440

Task 441

Task 442

Task 443

Task 444

Task 445

Task 446

Task 447

Task 448

Task 449

Task 450

Task 451

Task 452

Task 453

Task 454

Task 455

Task 456

Task 457

Task 458

Task 459

Task 460

Task 461

Task 462

Task 463

Task 464

Task 465

Task 466

Task 467

Task 468

Task 469

Task 470

Task 471

Task 472

Task 473

Task 474

Task 475

Task 476

Task 477

Task 478

Task 479

Task 480

Task 481

Task 482

Task 483

Task 484

Task 485

Task 486

Task 487

Task 488

Task 489

Task 490

Task 491

Task 492

Task 493

Task 494

Task 495

Task 496

Task 497

Task 498

Task 499

Task 500

Task 501

Task 502

Task 503

Task 504

Task 505

Task 506

Task 507

Task 508

Task 509

Task 510

Task 511

Task 512

Task 513

Task 514

Task 515

Task 516

Task 517

Task 518

Task 519

Task 520

Task 521

Task 522

Task 523

Task 524

Task 525

Task 526

Task 527

Task 528

Task 529

Task 530

Task 531

Task 532

Task 533

Task 534

Task 535

Task 536

Task 537

Task 538

Task 539

Task 540

Task 541

Task 542

Task 543

Task 544

Task 545

Task 546

Task 547

Task 548

Task 549

Task 550

Task 551

Task 552

Task 553

Task 554

Task 555

Task 556

Task 557

Task 558

Task 559

Task 560

Task 561

Task 562

Task 563

Task 564

Task 565

Task 566

Task 567

Task 568

Task 569

Task 570

Task 571

Task 572

Task 573

Task 574

Task 575

Task 576

Task 577

Task 578

Task 579

Task 580

Task 581

Task 582

Task 583

Task 584

Task 585

Task 586

Task 587

Task 588

Task 589

Task 590

Task 591

Task 592

Task 593

Task 594

Task 595

Task 596

Task 597

Task 598

Task 599

Task 600

Task 601

Task 602

Task 603

Task 604

Task 605

Task 606

Task 607

Task 608

Task 609

Task 610

Task 611

Task 612

Task 613

Task 614

Task 615

Task 616

Task 617

Task 618

Task 619

Task 620

Task 621

Task 622

Task 623

Task 624

Task 625

Task 626

Task 627

Task 628

Task 629

Task 630

Task 631

Task 632

Task 633

Task 634

Task 635

Task 636

Task 637

Task 638

Task 639

Task 640

Task 641

Task 642

Task 643

Task 644

Task 645

Task 646

Task 647

Task 648

Task 649

Task 650

Task 651

Task 652

Task 653

Task 654

Task 655

Task 656

Task 657

Task 658

Task 659

Task 660

Task 661

Task 662

Task 663

Task 664

Task 665

Task 666

Task 667

Task 668

Task 669

Task 670

Task 671

Task 672

Task 673

Task 674

Task 675

Task 676

Task 677

Task 678

Task 679

Task 680

Task 681

Task 682

Task 683

Task 684

Task 685

Task 686

Task 687

Task 688

Task 689

Task 690

Task 691

Task 692

Task 693

Task 694

Task 695

Task 696

Task 697

Task 698

Task 699

Task 700

Task 701

Task 702

Task 703

Task 704

Task 705

Task 706

Task 707

Task 708

Task 709

Task 710

Task 711

Task 712

Task 713

Task 714

Task 715

Task 716

Task 717

Task 718

Task 719

Task 720

Task 721

Task 722

Task 723

Task 724

Task 725

Task 726

Task 727

Task 728

Task 729

Task 730

Task 731

Task 732

Task 733

Task 734

Task 735

Task 736

Task 737

Task 738

Task 739

Task 740

Task 741

Task 742

Task 743

Task 744

Task 745

Task 746

Task 747

Task 748

Task 749

Task 750

Task 751

Task 752

Task 753

Task 754

Task 755

Task 756

Task 757

Task 758

Task 759

Task 760

Task 761

Task 762

Task 763

Task 764

Task 765

Task 766

Task 767

Task 768

Task 769

Task 770

Task 771

Task 772

Task 773

Task 774

Task 775

Task 776

Task 777

Task 778

Task 779

Task 780

Task 781

Task 782

Task 783

Task 784

Task 785

Task 786

Task 787

Task 788

Task 789

Task 790

Task 791

Task 792

Task 793

Task 794

Task 795

Task 796

Task 797

Task 798

Task 799

Task 800

Task 801

Task 802

Task 803

Task 804

Task 805

Task 806

Task 807

Task 808

Task 809

Task 810

Task 811

Task 812

Task 813

Task 814

Task 815

Task 816

Task 817

Task 818

Task 819

Task 820

Task 821

Task 822

Task 823

Task 824

Task 825

Task 826

Task 827

Task 828

Task 829

Task 830

Task 831

Task 832

Task 833

Task 834

Task 835

Task 836

Task 837

Task 838

Task 839

Task 840

Task 841

Task 842

Task 843

Task 844

Task 845

Task 846

Task 847

Task 848

Task 849

Task 850

Task 851

Task 852

Task 853

Task 854

Task 855

Task 856

Task 857

Task 858

Task 859

Task 860

Task 861

Task 862

Task 863

Task 864

Task 865

Task 866

Task 867

Task 868

Task 869

Task 870

Task 871

Task 872

Task 873

Task 874

Task 875

Task 876

Task 877

Task 878

Task 879

Task 880

Task 881

Task 882

Task 883

Task 884

Task 885

Task 886

Task 887

Task 888

Task 889

Task 890

Task 891

Task 892

Task 893

Task 894

Task 895

Task 896

Task 897

Task 898

Task 899

Task 900

Task 901

Task 902

Task 903

Task 904

Task 905

Task 906

Task 907

Task 908

Task 909

Task 910

Task 911

Task 912

Task 913

Task 914

Task 915

Task 916

Task 917

Task 918

Task 919

Task 920

Task 921

Task 922

Task 923

Task 924

Task 925

Task 926

Task 927

Task 928

Task 929

Task 930

Task 931

Task 932

Task 933

Task 934

Task 935

Task 936

Task 937

Task 938

Task 939

Task 940

Task 941

Task 942

Task 943

Task 944

Task 945

Task 946

Task 947

Task 948

Task 949

Task 950

Task 951

Task 952

Task 953

Task 954

Task 955

Task 956

Task 957

Task 958

Task 959

Task 960

Task 961

Task 962

Task 963

Task 964

Task 965

Task 966

Task 967

Task 968

Task 969

Task 970

Task 971

Task 972

Task 973

Task 974

Task 975

Task 976

Task 977

Task 978

Task 979

Task 980

Task 981

Task 982

Task 983

Task 984

Task 985

Task 986

Task 987

Task 988

Task 989

Task 990

Task 991

Task 992

Task 993

Task 994

Task 995

Task 996

Task 997

Task 998

Task 999

Task 1000

Assignment5-task1 - Spark Jobs - Google Chrome

Assign x My Cla x Workb x EMR - x Datap x EMR - x Assign x 10186 x 1.4.4 - x java - x LogMi x 15.7. lc x Python x +

Not secure | ec2-34-204-167-23.compute-1.amazonaws.com:18080/history/application_1587164871837_0001/1/jobs/

Spark 2.4.4 Jobs Stages Storage Environment Executors

Assignment5-task1 application UI

Spark Jobs (?)

User: hadoop
Total Uptime: 1.5 h
Scheduling Mode: FIFO
Completed Jobs: 79

Event Timeline
Enable zooming

Executors
Added
Removed
Executor driver added
Executor 1 added

Jobs
Succeeded
Failed
Running

runJob at PythonRDD scala:153 (Job 1)

treeAggreg

top at max

collect

23:10
Fri 17 April

23:15

23:20

23:25

23:30

23:35

23:40

23:45

23:50

23:55

00:00
Sat 18 April

00:05

00:10

00:15

00:20

00:25

00:30

00:35

00:40

00:45

Completed Jobs (79)

Job Id	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
78	runJob at SparkHadoopWriter.scala:78 runJob at SparkHadoopWriter.scala:78	2020/04/18 00:42:54	0.4 s	1/1	1/1

S3 Management Console - Google Chrome

Assign x My Cla x Workb x EMR- x Datap x S3 Mar x Assign x 10186 x 1.4.4- x java- x LogM x 15.7.1c x Python x +

s3.console.aws.amazon.com/s3/buckets/gagankaushal/Assignment_5/?region=us-east-1&tab=overview

aws Services Resource Groups

vocstartsoft/user593296-gk@bu.e... Global Support

Amazon S3 > gagankaushal > Assignment_5

gagankaushal

Overview

Type a prefix and press Enter to search. Press ESC to clear.

Upload Create folder Download Actions

US East (N. Virginia)

Viewing 1 to 2

<input type="checkbox"/> Name	Last modified	Size	Storage class
<input type="checkbox"/> Output_task1	--	--	--
<input type="checkbox"/> main_task1.py	Apr 17, 2020 2:03:13 PM GMT-0400	12.6 KB	Standard

Viewing 1 to 2

Feedback English (US) © 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

S3 Management Console - Google Chrome

Assign x My Cla x Workb x EMR- x Datap x S3 Mar x Assign x 10186 x 1.4.4- x java- x LogM x 15.7.1c x Python x +

s3.console.aws.amazon.com/s3/buckets/gagankaushal/Assignment_5/?region=us-east-1&tab=overview

aws Services Resource Groups

vocstartsoft/user593296-gk@bu.e... Global Support

Amazon S3 > gagankaushal > Assignment_5

gagankaushal

Overview

Type a prefix and press Enter to search. Press ESC to clear.

Upload Create folder Download Actions

US East (N. Virginia)

Viewing 1 to 2

<input type="checkbox"/> Name	Last modified	Size	Storage class
<input type="checkbox"/> Output_task1	--	--	--
<input type="checkbox"/> main_task1.py	Apr 17, 2020 2:03:13 PM GMT-0400	12.6 KB	Standard

Viewing 1 to 2

Feedback English (US) © 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Assignment_5_Output_task1_part-00000 (~/Desktop/BU/Spring2020/Big Dat...Analytics/Assignment5/Final/docs) - gedit

Assignment_5_Output_task1_part-00000
~/Desktop/BU/Spring2020/Big Dat...Analytics/Assignment5/Final/docs

Save

```
('F1 score for Logistic Regression classifier (trained using mllib)' =', 99.6, '%')  
  
('Time taken to read Testing and Training Data and preprocess them (days:seconds:microseconds):', datetime.timedelta(0, 230, 238702))  
('Time taken to train the Logistic Regression model (days:seconds:microseconds):', datetime.timedelta(0, 5167, 335214))  
('Time taken to test the Logistic Regression model (days:seconds:microseconds):', datetime.timedelta(0, 175, 346919))  
('Total Time taken by Logistic Regression (trained using mllib) (days:seconds:microseconds):', datetime.timedelta(0, 5572, 920835))  
  
('Number of True Positives', 374)  
('Number of False Positives', 0)  
('Number of False Negatives', 3)  
('Number of True Negatives', 18347)
```

Plain Text ▾ Tab Width: 8 ▾ Ln 13, Col 1 ▾ INS

P. T.O.

Task 2

App ID: [application_1587174763389_0002](#)

https://kbykz5pprvbodbkfz4hhqc2r6u-dot-us-east1.dataproc.googleusercontent.com/sparkhistory/ - Google Chrome

Assignn x Clusters x https:// x https:// x 101867 x met-cs-7 x metcs/r x EMR-A x Assignn x EMR-A x Assignn x Assignn x +

← → ↻ kbykz5pprvbodbkfz4hhqc2r6u-dot-us-east1.dataproc.googleusercontent.com/sparkhistory/ ☆ 6 ⋮

evident-hexagon-113810 > assignment5-2 Sign out

Spark 2.3.4 **History Server**

Event log directory: hdfs://assignment5-2-m/user/spark/eventlog
Last updated: 2020-04-18 00:38:23
Client local time zone: America/New_York

Search:

App ID	App Name	Started	Completed	Duration	Spark User	Last Updated	Event Log
application_1587174763389_0002	Assignment5-task2	2020-04-17 22:38:14	2020-04-18 00:28:30	1.8 h	root	2020-04-18 00:28:30	Download
application_1587174763389_0001	Assignment5-task2	2020-04-17 21:57:26	2020-04-17 22:37:26	40 min	root	2020-04-17 22:37:26	Download

Showing 1 to 2 of 2 entries
[Show incomplete applications](#)

https://kbykz5pprvbodbkfz4hhqc2r6u-dot-us-east1.dataproc.googleusercontent.com/gateway/default/sparkhistory/history/application_1587174763389_0002/jobs/ - Google Chrome

Assignn x Clusters x https:// x https:// x 101867 x met-cs-7 x metcs/r x EMR-A x Assignn x EMR-A x Assignn x Assignn x +

← → ↻ kbykz5pprvbodbkfz4hhqc2r6u-dot-us-east1.dataproc.googleusercontent.com/gateway/default/sparkhistory/history/application_1587174763389_0002/jobs/ 🔍 ☆ 6 ⋮

evident-hexagon-113810 > assignment5-2 Sign out

Spark 2.3.4 Jobs Stages Storage Environment Executors Assignment5-task2 application UI

Spark Jobs (?)

User: root
Total Uptime: 1.8 h
Scheduling Mode: FAIR
Completed Jobs: 517

▼ Event Timeline
☑ Enable zooming

Executors
Added
Removed

Jobs
Succeeded
Failed
Running

top all
count at hlogassignment5-task2_withcache_32_3amd

02:40 02:45 02:50 02:55 03:00 03:05 03:10 03:15 03:20 03:25 03:30 03:35 03:40 03:45 03:50 03:55 04:00 04:05 04:10 04:15 04:20 04:25 04:30 0

Sat 18 April

Completed Jobs (517)

Page: 1 2 3 4 5 6 >

Job Id	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
516	runJob at SparkHadoopWriter.scala:78 runJob at SparkHadoopWriter.scala:78	2020/04/18 04:28:27	2 s	1/1	1/1
515	runJob at SparkHadoopWriter.scala:78 runJob at SparkHadoopWriter.scala:78	2020/04/18 04:28:24	2 s	1/1	1/1

6 Pages. Jump to 1 . Show 100 items in a page. Go

Bucket details - My Project - Google Cloud Platform - Google Chrome

Assignn x Bucket c x https:// x 101867 x met-cs-7 x metcs/n x EMR-A x Assignn x EMR-A x Assignn x Assignn x +

console.cloud.google.com/storage/browser/gagankaushal.com/Assignment_5/Final_/?forceOnBucketsSortingFiltering=false&authuser=0&project=evident-hexagon-113810

Your free trial is waiting: activate now to get \$300 credit to explore Google Cloud products. [Learn more](#) DISMISS ACTIVATE

Google Cloud Platform My Project Search resources and products

Storage

Browser

Transfer

Transfer for on-premises

Transfer Appliance

Settings

Bucket details EDIT BUCKET REFRESH BUCKET

gagankaushal.com

Objects Overview Permissions Bucket Lock

Upload files Upload folder Create folder Manage holds Delete

Filter by prefix...

Buckets / gagankaushal.com / Assignment_5 / Final_

<input type="checkbox"/>	Name	Size	Type	Storage class	Last modified	Public access	Encryption	Retention expiration date	Holds
<input type="checkbox"/>	Output_task2/	-	Folder	-	-	Subject to object ACLs	-	-	-
<input type="checkbox"/>	main_task2.py	15.38 KB	text/x-python	Standard	4/17/20, 10:34:47 PM UTC-4	Not public	Google-managed key	-	None

Bucket details - My Project - Google Cloud Platform - Google Chrome

Assignn x Bucket c x https:// x https:// x 101867 x met-cs-7 x metcs/n x EMR-A x Assignn x EMR-A x Assignn x Assignn x +

console.cloud.google.com/storage/browser/gagankaushal.com/Assignment_5/Final_/_Output_task2/?forceOnBucketsSortingFiltering=false&authuser=0&project=evident-hexagon-113810

Your free trial is waiting: activate now to get \$300 credit to explore Google Cloud products. [Learn more](#) DISMISS ACTIVATE

Google Cloud Platform My Project Search resources and products

Storage

Browser

Transfer

Transfer for on-premises

Transfer Appliance

Settings

Bucket details EDIT BUCKET REFRESH BUCKET

gagankaushal.com

Objects Overview Permissions Bucket Lock

Upload files Upload folder Create folder Manage holds Delete

Filter by prefix...

Buckets / gagankaushal.com / Assignment_5 / Final_ / Output_task2

<input type="checkbox"/>	Name	Size	Type	Storage class	Last modified	Public access	Encryption	Retention expiration date	Holds
<input type="checkbox"/>	_SUCCESS	0 B	application/octet-stream	Standard	4/18/20, 12:28:27 AM UTC-4	Not public	Google-managed key	-	None
<input type="checkbox"/>	part-00000	594 B	application/octet-stream	Standard	4/18/20, 12:28:26 AM UTC-4	Not public	Google-managed key	-	None

Assignment_5_Final_Output_task2_part-00000 (-/Desktop/BU/Spring2020/Big ...ytics/Assignment5/Final/docs) - gedit

Assignment_5_Final_Output_task2_part-00000
~/Desktop/BU/Spring2020/Big ...ytics/Assignment5/Final/docs

Open Save

*Untitled Document 1 main_task3.py next_steps.txt main_task2.py Assignment_5_Output_task1_part-00000 Assignment_5_Final_...put_task2_part-00000

```
('F1 score for classifier =', 99.87, '%')

('Time taken to read Testing and Training Data and preprocess them (days:seconds:microsecond):', datetime.timedelta(0, 1292, 854187))
('Time taken to train using the Support Vector Machine (days:seconds:microsecond):', datetime.timedelta(0, 5148, 112811))
('Time taken to test using the Support Vector Machine (days:seconds:microsecond):', datetime.timedelta(0, 158, 967617))
('Total Time taken by Support Vector Machine (days:seconds:microsecond):', datetime.timedelta(0, 6599, 934615))

('Number of True Positives', 377)
('Number of False Positives', 1)
('Number of False Negatives', 0)
('Number of True Negatives', 18346)
```

Plain Text Tab Width: 8 Ln 3, Col 94 INS

P.T.O.

Task 3

App ID: [application_1587164871837_0003](#)

History Server - Google Chrome

1018: Data | metc | EMR | Assig | S3 M | Assig | HI x | Assig | Assig | Assig | Histc | Assig | Assig | Assig | Assig | Assig | met- | Assig | +

Not secure | ec2-34-204-167-23.compute-1.amazonaws.com:18080

History Server

Event log directory: `hdfs://var/log/spark/apps`
Last updated: 2020-04-18 05:28:38
Client local time zone: America/New_York

Search:

App ID	App Name	Started	Completed	Duration	Spark User	Last Updated	Event Log
application_1587164871837_0003	Assignment5-task3	2020-04-17 21:45:37	2020-04-18 02:20:38	2.3 h	hadoop	2020-04-18 02:20:38	Download
application_1587164871837_0002	Assignment5-task3	2020-04-17 21:24:00	2020-04-17 21:58:53	35 min	hadoop	2020-04-17 21:58:53	Download
application_1587164871837_0001	Assignment5-task1	2020-04-17 19:09:58	2020-04-17 20:42:54	1.5 h	hadoop	2020-04-17 20:42:54	Download

Showing 1 to 3 of 3 entries
[Show incomplete applications](#)

Assignment5-task3 - Spark Jobs - Google Chrome

1018: Data | metc | EMR | Assig | S3 M | Assig | As x | Assig | Assig | Assig | Histc | Assig | Assig | Assig | Assig | Assig | met- | Assig | +

Not secure | ec2-34-204-167-23.compute-1.amazonaws.com:18080/history/application_1587164871837_0003/1/jobs/

Spark Jobs (?)

User: hadoop
Total Uptime: 2.3 h
Scheduling Mode: FIFO
Completed Jobs: 429

Event Timeline
☐ Enable zooming

Executors
☐ Added
☐ Removed

Jobs
☐ Succeeded
☐ Failed
☐ Running

▼ Completed Jobs (429)

Page: 1 2 3 4 5 >

Job id	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
428	runJob at SparkHadoopWriter.scala:78	2020/04/18 06:20:37	0.3 s	1/1	1/1


```
Assignment_5_Final_Output_task3_part-00000
~/Desktop/BU/Spring2020/Big ...n on AWS although GCP passed Save

('#####', ' LOGISTIC REGRESSION - RESULTS ', '#####')
('F1 score for classifier =', 99.6, '%')

('Time taken to read Testing and Training Data and preprocess them (trained using mlb)(days:seconds:microsecond:', datetime.timedelta(0, 2400, 746235))
('Time taken to train the Logistic Regression model (trained using mlb)(days:seconds:microsecond:', datetime.timedelta(0, 1322, 621003))
('Time taken to test the Logistic Regression model (trained using mlb)(days:seconds:microsecond:', datetime.timedelta(0, 156, 601166))
('Total Time taken by Logistic Regression (trained using mlb)(days:seconds:microsecond:', datetime.timedelta(0, 3879, 968404))

('Number of True Positives', 374)
('Number of False Positives', 0)
('Number of False Negatives', 3)
('Number of True Negatives', 18347)

('#####', ' SUPPORT VECTOR - RESULTS ', '#####')
('F1 score for classifier =', 98.79, '%')

('Time taken to read Testing and Training Data and preprocess them (days:seconds:microsecond):', datetime.timedelta(0, 2400, 746235))
('Time taken to train the Support Vector Machine model(days:seconds:microsecond):', datetime.timedelta(0, 1415, 655597))
('Time taken to test using the Support Vector Machine model (days:seconds:microsecond):', datetime.timedelta(0, 138, 874487))
('Total Time taken by Support Vector Machine (days:seconds:microsecond):', datetime.timedelta(0, 3955, 276319))

('Number of True Positives', 368)
('Number of False Positives', 0)
('Number of False Negatives', 9)
('Number of True Negatives', 18347)
```