

# Experiment Management and Collaboration

Lesson 6

---

DVC Tools for Data Scientists &  
Analysts

2021



# Course lessons

**Lesson 1.** Course Introduction

**Lesson 2.** Practices and tools for Efficient Collaboration in ML projects

**Lesson 3.** Pipelines Automation and Configuration Management

**Lesson 4.** Versioning Data and Models

**Lesson 5.** Visualize Metrics & Compare Experiments with DVC and Studio

**Lesson 6.** Experiments Management and Collaboration

**Lesson 7.** Tools for Deep Learning Scenarios

**Lesson 8.** Review Advanced Topics and Use Cases



## Lesson Outline

- ◇ Experiment management with **dvc exp**
- ◇ Experimenting workflow with Git, DVC and Studio





# Experiment management

**with** `dvc exp run`

# Experiment management with **dvc exp**



- ◆ Runs DVC experiments (similar to **dvc repro**)
- ◆ Experiment specific features
  - update params
  - queueing experiments
  - comparing many experiments
  - persisting experiments
  - sharing experiments

```
$ dvc exp show --no-timestamp \  
--include-params train.n_est,train.min_split
```

Experiment	avg_prec	roc_auc	train.n_est	train.min_split
workspace	0.56191	0.93345	50	2
master	0.55259	0.91536	50	2
└─ exp-bfe64	0.57833	0.95555	50	8
└─ exp-b8082	0.59806	0.95287	50	64
└─ exp-c7250	0.58876	0.94524	100	2
└─ exp-b9cd4	0.57953	0.95732	100	8
└─ exp-98a96	0.60405	0.9608	100	64
└─ exp-ad5b1	0.56191	0.93345	50	2

# Modify parameters on-the-fly



# run command

```
$ dvc exp run -S prepare.split=0.25 -S featurize.max_features=2000
```

The **--set-param** (or **-S**) flag sets the values for parameters as a shortcut to editing **params.yaml**

# Queueing experiments



# run command

```
$ dvc exp run --queue -S featurize.max_features=20  
$ dvc exp run --queue -S featurize.max_features=1000  
$ dvc exp run --queue -S featurize.max_features=2000
```

Setup experiments with  
different hyperparameters  
with **--queue** flag

# run command

```
$ dvc exp run --run-all --jobs 2
```

Run all queued  
experiments using  
**--run-all** (and in parallel  
with **--jobs**)

# Navigate experiments and metrics in CLI



# run command

```
$ dvc exp show --include-params=featurize
```

Experiment	Created	auc	featurize.max_features
workspace	–	0.61314	1500
10-bigrams-experiment	Jun 20, 2020	0.61314	1500
└─ exp-e6c97	Oct 21, 2020	0.61314	1500
└─ exp-1dad0	Oct 09, 2020	0.57756	2000
└─ exp-1df77	Oct 09, 2020	0.51676	500



# Make an experiment persistent



# run command

```
$ dvc exp apply exp-e6c97
```

Experiment	Created	auc	featurize.max_features
workspace	–	0.61314	1500
10-bigrams-experiment	Jun 20, 2020	0.61314	1500
├─ exp-e6c97	Oct 21, 2020	0.61314	1500
├─ exp-10ad0	Oct 09, 2020	0.57756	2000
└─ exp-10f77	Oct 09, 2020	0.51676	

Changes for experiment  
'exp-e6c97' have been  
added to the workspace

# Make a persistent branch from an experiment



# run command

```
$ dvc exp branch exp-ldad0 maxf-2000
```

Git branch 'maxf-2000' has been created from experiment 'exp-ldad0'.

To switch to the new branch run:  
git checkout my-branch

**maxf-2000** can now be  
checked out, merged,  
rebased, pushed, etc. like  
any other Git branch



Live code example

# Experiment Management with `dvc exp run`

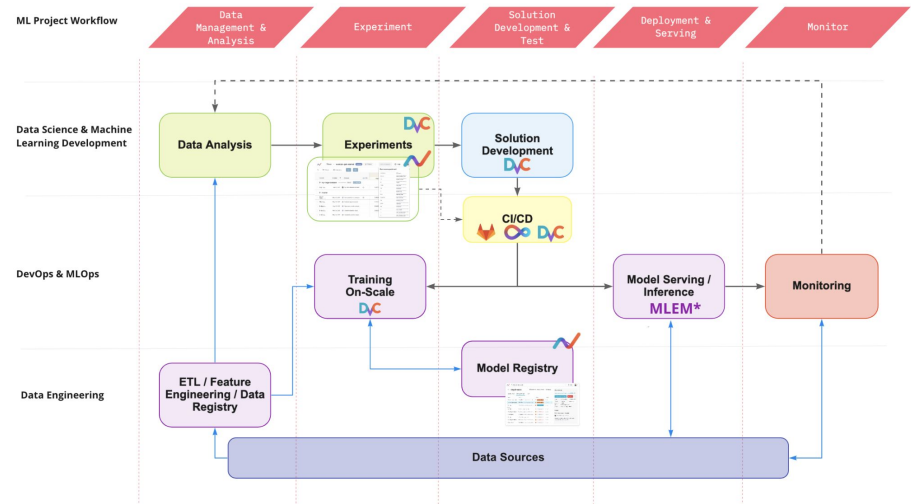


# Experimenting Workflow with DVC and Studio

# Experimenting Workflow



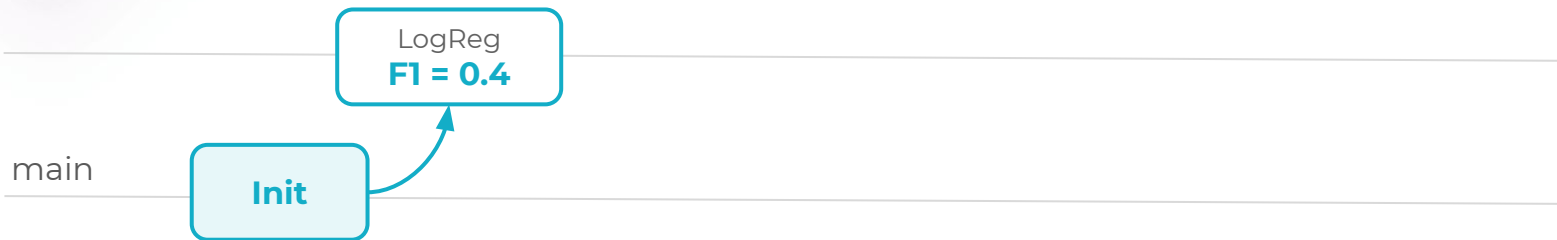
- ◇ **Git:** branching and tagging experiments
  - a. create a separate branch & tag for each experiment
  - b. keep failed experiments
- ◇ **DVC:** run pipelines and keep states
  - a. use **dvc exp run**
  - b. clean up DVC cache for not important experiments
- ◇ **Studio:** visualize metrics and collaborate
  - a. metrics visualization
  - b. experiments comparison
  - c. collaboration



# Run experiment: `dvc exp run`



Experiment	avg_prec	roc_auc	train.n_est	train.min_split
workspace	0.56191	0.93345	50	2
master	0.55259	0.91536	50	2
└ exp-bfe64	0.57833	0.95555	50	8
└ exp-ad5b1	0.56191	0.93345	50	2



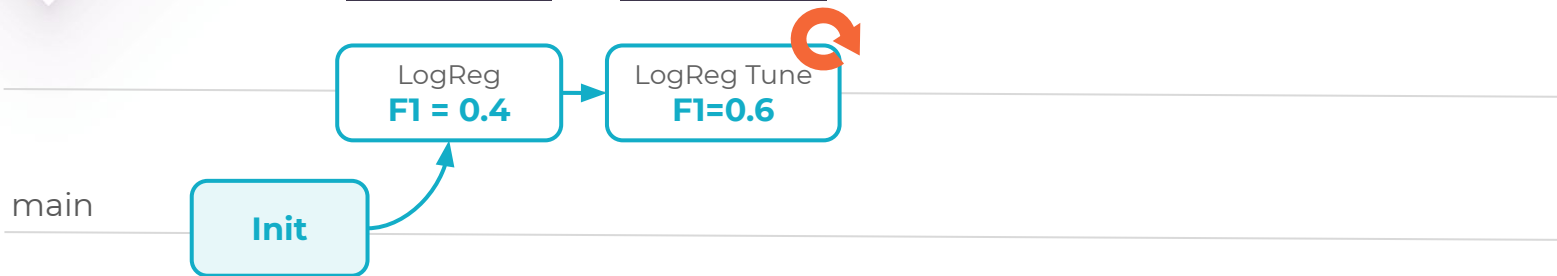
`dvc exp run` to run experiment  
`dvc exp show` to track metrics in CLI

# Update code/params, continue: `dvc exp run`



Experiment	Model	Model	Model	Model
LogReg	0.5015	0.5015	0.5015	0.5015
LogReg	0.5015	0.5015	0.5015	0.5015
LogReg	0.5015	0.5015	0.5015	0.5015
LogReg	0.5015	0.5015	0.5015	0.5015

Experiment	Model	Model	Model	Model
LogReg	0.5015	0.5015	0.5015	0.5015
LogReg	0.5015	0.5015	0.5015	0.5015
LogReg	0.5015	0.5015	0.5015	0.5015
LogReg	0.5015	0.5015	0.5015	0.5015



Update code/params  
`dvc exp run` to run experiment  
`dvc exp show` to track metrics in CLI

# Commit successful experiment



experiment 1

Experiment	trial_name	log_prob	train_loss	train_acc
experiment	0.10131	0.00105	0.0	0.0
experiment	0.10131	0.00105	0.0	0.0
experiment	0.10131	0.00105	0.0	0.0
experiment	0.10131	0.00105	0.0	0.0



LogReg  
F1 = 0.4

LogReg Tune  
F1=0.6

main

Init

option 1 (on the slide): **create a new branch**

**dvc exp branch** <exp> **experiment-1**

**dvc push** tracked files to remote storage  
**git commit & push** to keep updates in Git

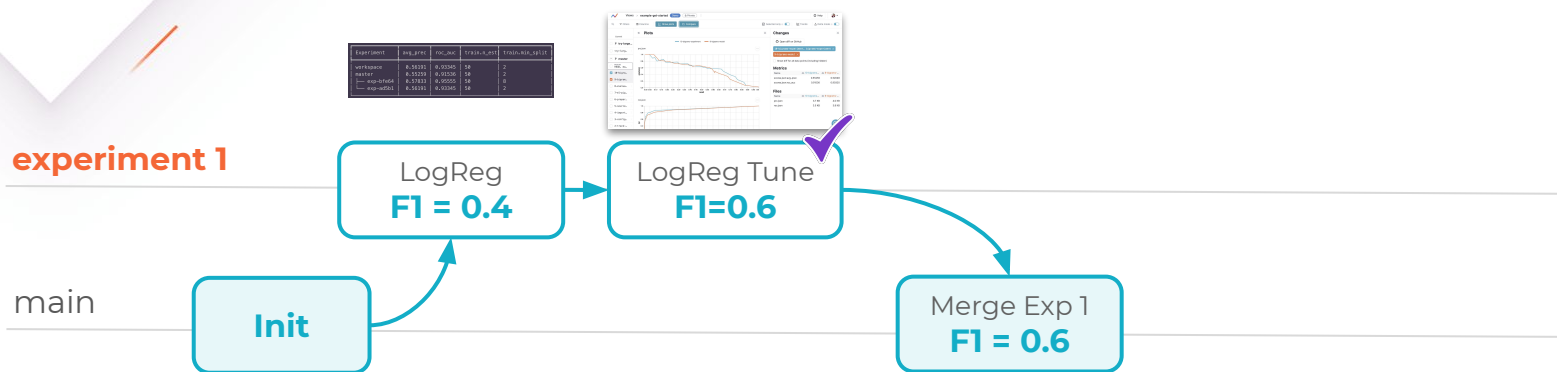
option 2: **apply to the current branch**

**dvc exp apply** <exp>

...



# Merge successful experiment to **main** branch



Merge **experiment-1** branch to **main**  
(using GitLab/GitHub/BitBucket interface)

Or keep it in a separate branch

# Continue experimenting



experiment 1

Experiment	MLP	Log_2	1212-A-101	1000-MLP
experiment	0.18131	0.02045	18	2
merge	0.15273	0.01505	18	2
exp-mlp1	0.17113	0.00505	18	9
exp-mlp1	0.18131	0.03045	18	7



LogReg  
**F1 = 0.4**

LogReg Tune  
**F1=0.6**

main

Init

Merge Exp 1  
**F1 = 0.6**

Merge SVM  
**F1=0.7**

experiment 2

SVM  
**F1=0.7**



Neural Nets  
**F1 = 0.5**

Tunned NN  
**F1=0.69**



## Tips:

- ◇ keep important experiments in history
- ◇ use Git tags if you need



Live code example

# Experimenting Workflow with DVC and Studio



# **Team Collaboration with DVC Studio**

# Studio: create a team workspace



**Personal** ▾ Views

- Personal ✓
- Teams
- Iterative
- CSE\_iterative
- mlrepa
- + Create a team

Find a view or repo + Add a view Talk to an expert

**mono-1-model-deps-solution-2**

2 issues Updated 2 months ago

iterative.ai/cse/use\_cases/mono-1-model-deps-solution-2

Local remote was ignored for /tmp/mo

Metrics

reports/model-c-metrics.json:f1_score	
reports/model-a-metrics.json:f1_score	0.94841
reports/model-b-metrics.json:f1_score	1.00000

**telecom-customer-churn**

1 issue

mnrozhkov/telecom-customer-churn

Metrics

lift_max	
lift_m	
lift_m	

**example-get-started**

Updated 5 days ago

iterative/example-get-started

Metrics

avg_prec	
roc_auc	

**Navigate or Create a Team workspace**

**Connect to Git repository and create a "View"**

**Make views public**

# Studio: explore colleagues contributions



Views > checkpoints-tutorial Private You have unsaved changes Discard Save

Local remote was ignored for /tmp/dvc/checkpoints-tutorials

Filters 1 Columns Show plots Compare Run

Clear Author is Mikhail X

+ Add filter

					dvclive.json			
					acc	loss	step	dvclive model.pt predictions.json
<input type="checkbox"/>	dev	inherited from main						
<input type="checkbox"/>	dev	Oct 29, 2021	Good experiment		0.96510	0.12553	33	1.3 KB 410.2 KB 310.0 KB
<input type="checkbox"/>	57de8db	Oct 29, 2021	Good experiment		0.96510	0.12553	33	• 1.3 KB • 410.2 KB • 310.0 KB
<input type="checkbox"/>	6e0c215	Oct 08, 2021	saved files from experiment		0.96350	0.13060	23	• 1.5 KB • 410.2 KB • 310.0 KB
<input type="checkbox"/>	f54e823	Oct 08, 2021	Create DVC pipeline		—	—	—	— — —

Filter experiments by Author name

# Studio: manage a team permissions



## Team

Adding a collaborator will give them access to all the **Iterative** views.

### Invite by email

Can edit ▾

Invite

### Collaborators (5 of 10)



Alexander Schepanovski

@gmail.com

Admin ▾



Ivan Shcheklein

@gmail.com

Admin ▾



Maxim Shmakov

@mail.ru

Admin ▾



tapadipti You

@gmail.com

Admin ▾



Mikhail

@gmail.com

Admin ▾



Manage  
collaborators and  
permissions



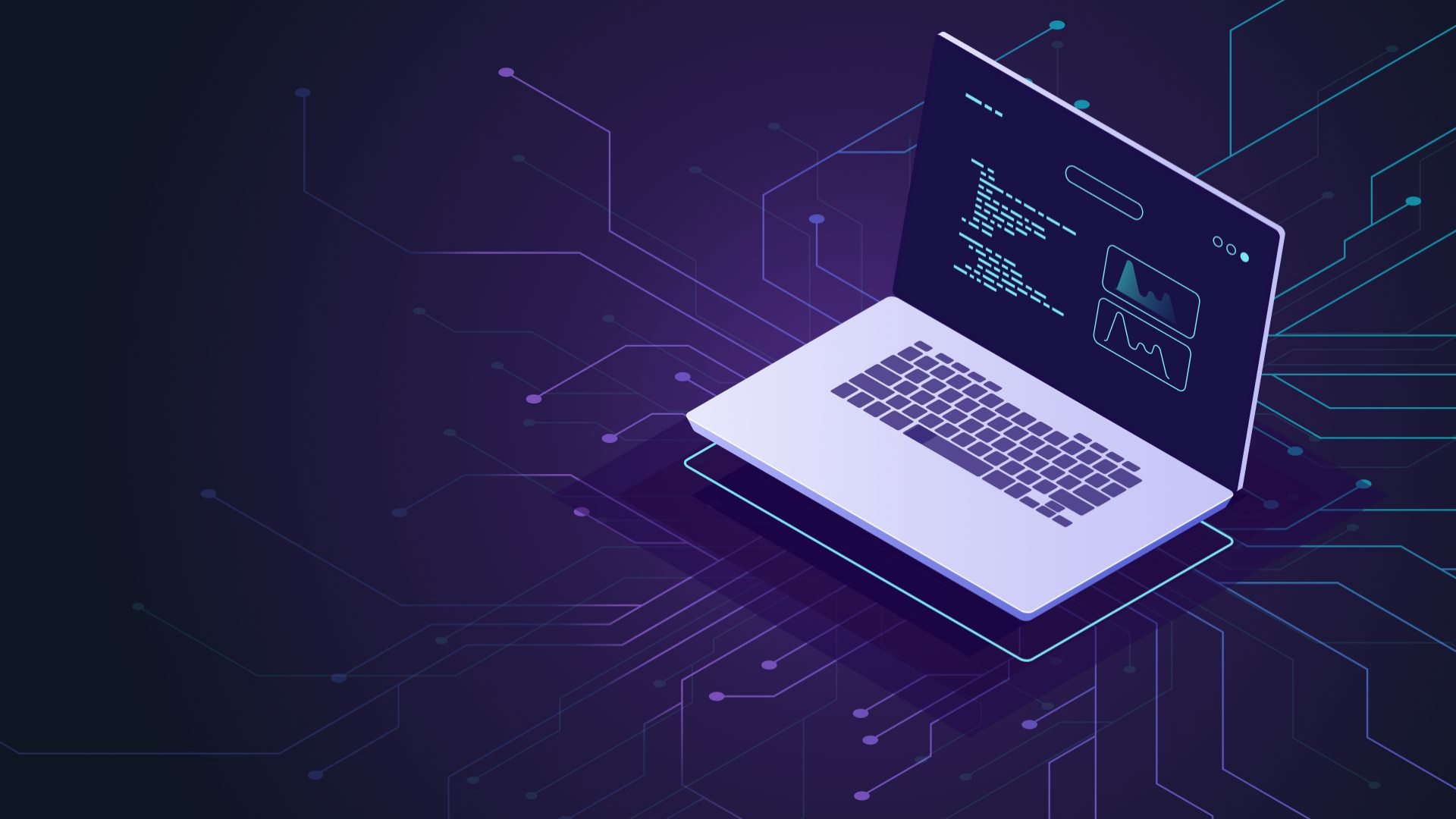
**What have we learned?**



# What have we learned?



1. **dvc exp run**
2. Experimenting workflow with Git, DVC and Studio
3. Team Collaboration with DVC Studio





# Links

- ◆ Experiment Management <https://dvc.org/doc/user-guide/experiment-management>
- ◆ Docs: dvc exp <https://dvc.org/doc/command-reference/exp>