Scalable Machine Learning with Ray

SWUNG TRANSFORM 2020 - Friday 12th June 9am (BST)

Steve Purves

swung: stevejpurves
twitter: @stevejpurves
github: stevejpurves
Email: stevejpurves@gmail.com

https://github.com/euclidity/transform-2020-ray

“There are many more possible [hyper-]parameter variations than I have time/money/patience to try”

-- someone, somewhere tuning an ML model

- You can pause the livestream!
- Join the slack channel #t20-fri-ray
- @Dan A (Dan Austin) is also in the channel and can help out
- If you have questions ask and we'll try and answer them after the break
- The tutorial was developed and tested on macos & ubuntu. Windows is an unknown factor.

Let's get setup!

`Tutorial Repo is at:`
`https://github.com/euclidity/transform-2020-ray`

1. Ray Dashboard
   a. Install Node.js
2. Build the dashboard from source
3. Build the Conda Environment

If you can't get the dashboard built, don't worry and move on. You'll lose a little bit of interactivity but it can still work.

When we access tensorboard via the dashboard you can just launch this locally from a terminal:

- `tensorboard --logdir ~/ray_results/named_result_folder/`
- `http://localhost:6006`

Aims

This tutorial aims to:

- Introduce Ray it's packages
- Take a close look at Raytune for Hyperparameter Optimisation
- Do this though examples relevant to our domain as much as possible
- Show that if you code is in the right shape, moving to Raytune is simple
- Leave you with notebooks as resources to start playing with a building out from

15% ML Practices

20% Scikit Learn

15% Pytorch

50% Ray & Raytune

Examples:

General:

- Breast Cancer Dataset
- MNIST

Domain:

- Well Lithofacies
- Fossil Images

Content & Running Order

- Review Setup
- Intro & Running Order
- Warm Up in SKLearn - What and Why?
- Getting Started with Ray
- Tuning for Real - Well Data
- Deep Trainables - PyTorch

BREAK 15 mins

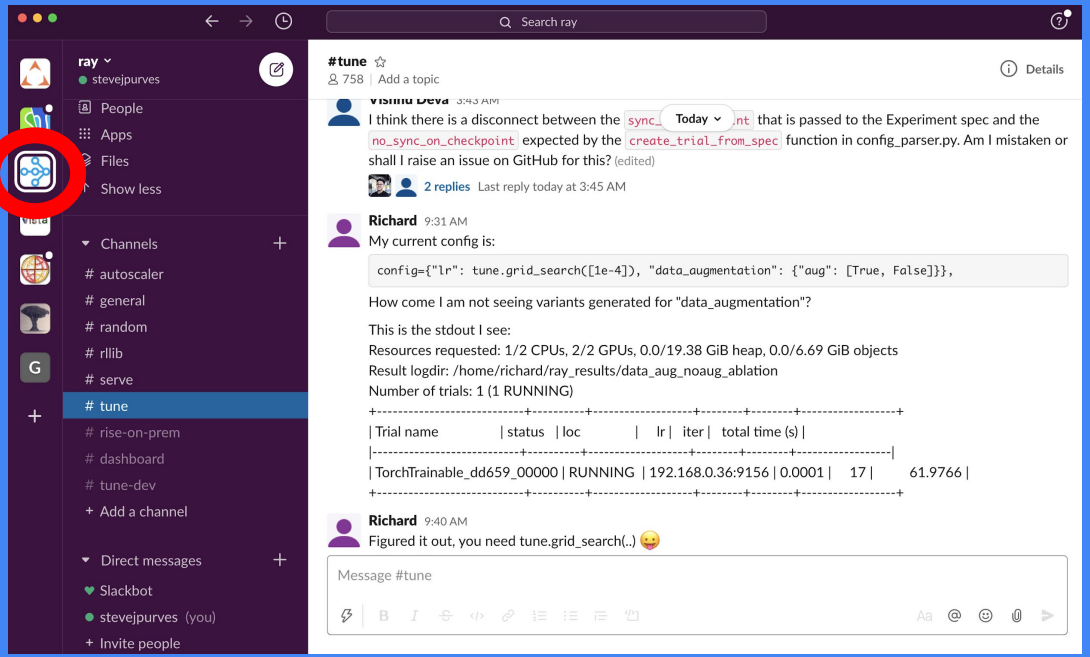
Let’s get started

What is Ray?

a distributed computation framework

infrastructure on a stick

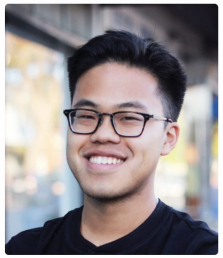
anyscale



Events

RAY SUMMIT CONNECT

Each month's *Ray Summit Connect* includes several short talks on the theme of the event, followed by a panel discussion with audience Q&A.



June 17, 2020, 10:00am

The Road to AutoML: Hyperparameter Tuning and Neural Architecture Search

The vision of AutoML is to remove as much manual effort and required expertise as possible when applying machine learning and artificial intelligence to real-world problems. This Ray Summit Connect explores two topics of...

More details

July 8, 2020, 10:00am

Practical Reinforcement Learning

The Ray Ecosystem

- Raytune
- RLLib [docs] [tutorials]
- RaySGD [docs]
- RayServe [pytorch tutorial]
- Ray sklearn [repo]

That’s all

Steve Purves

swung: stevejpurves
twitter: @stevejpurves
github: stevejpurves
Email: stevejpurves@gmail.com