

**Carnegie
Mellon
University**



ML-Ops with ClearML

for Movie Rating Predictions

10TH NOVEMBER 2022

Davidson Siga, Graduate Student @ CMU
For the ML in Production Class , CMU (17-645)

Agenda

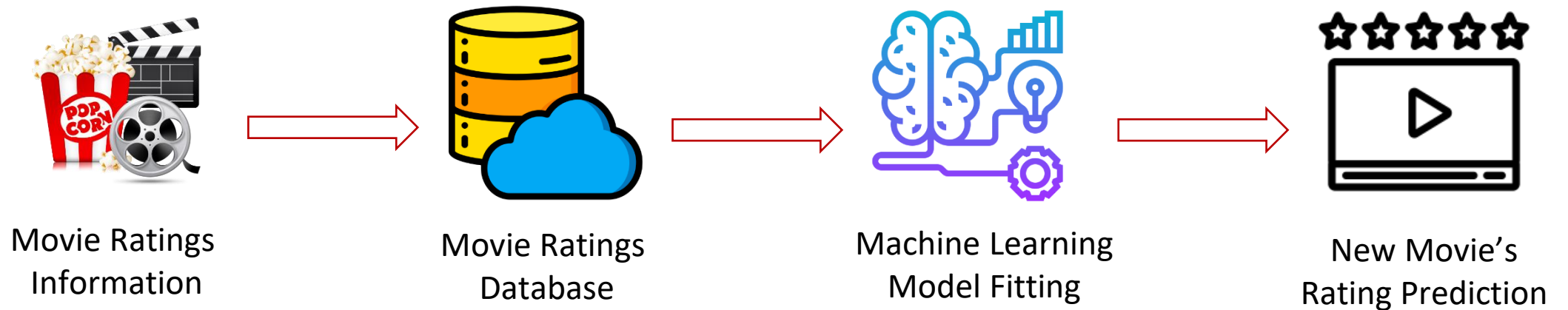
1. An ML Problem : Movie Rating Prediction
2. ClearML : ML-Ops for efficient productionizing
3. Integration : ClearML + Movie Rating Prediction
4. The Strengths and Limitations of the tool



An ML Problem

Movie Rating Prediction

Predicting Movie Ratings



Idea of Machine Learning Prediction:

Learn on data which looks like : User 1 watches Movie 1 and rates it 4/5 stars

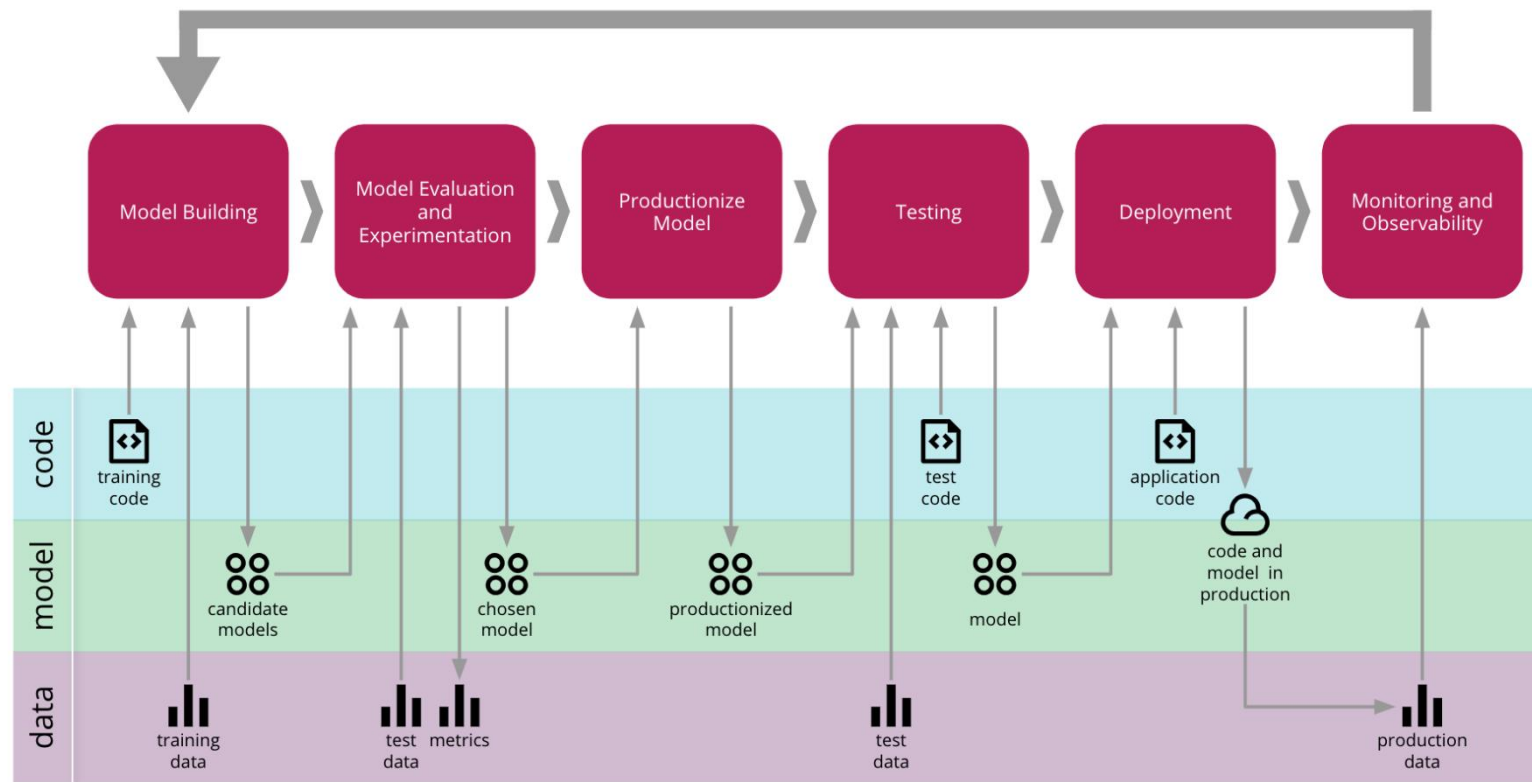
If we give model the Features : Demographics of User and Movie Details

It will give us the Outcome : Prediction (1-5 Stars)

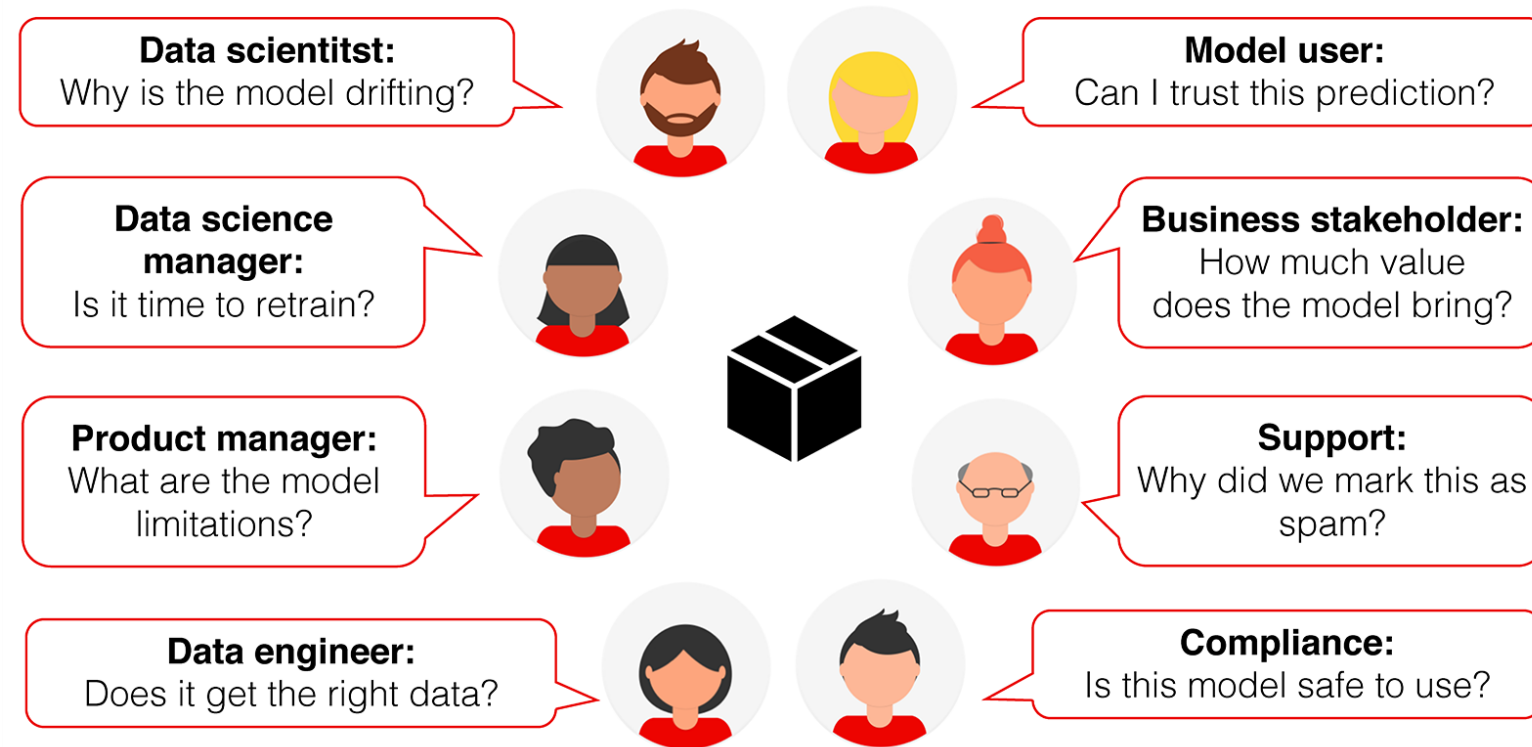
Looking at the Data In hand

	MovieIDs	MovieName	Category	ID	MovieID	Ratings	TimeStamp	UserID	Gender	Age	Occupation	Zip-code	age_group
0	2487.0	Blood, Guts, Bullets and Octane (1998)	Action Comedy	19	3081	4	978555317	2419.0	M	25.0	0.0	06096	20-29
1	2325.0	Orgazmo (1997)	Comedy	19	3863	4	978554522	2257.0	F	18.0	4.0	59715	10-19
2	3489.0	Hook (1991)	Adventure Fantasy	26	592	3	978146876	3421.0	M	45.0	7.0	94501	40-49
3	3042.0	Meatballs III (1987)	Comedy	23	3700	4	978464211	2974.0	M	18.0	4.0	48105	10-19
4	2250.0	Men Don't Leave (1990)	Drama	18	2014	2	978155339	2182.0	F	25.0	4.0	60443	20-29
...
3488	2231.0	Rounders (1998)	Crime Drama	18	2872	5	978152515	2163.0	M	18.0	0.0	87122	10-19
3489	3474.0	Retroactive (1997)	Sci-Fi Thriller	26	1	3	978130703	3406.0	F	25.0	7.0	85214	20-29
3490	64.0	Two if by Sea (1996)	Comedy Romance	2	3468	5	978298542	64.0	M	18.0	1.0	53706	10-19
3491	1483.0	Crash (1996)	Drama Thriller	13	2968	3	978202417	1454.0	M	18.0	4.0	93117	10-19

ML Engineering in General



Many Stakeholder, Many Coders = Chaos!



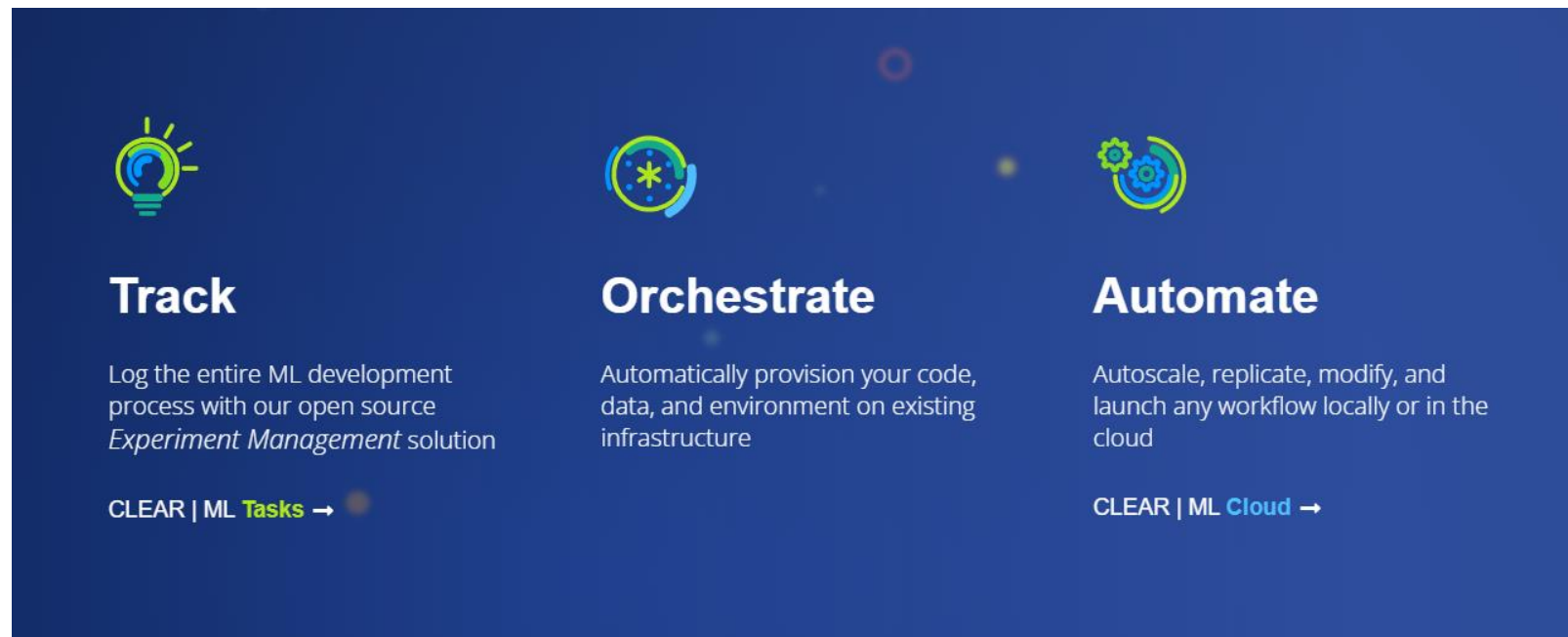


ClearML

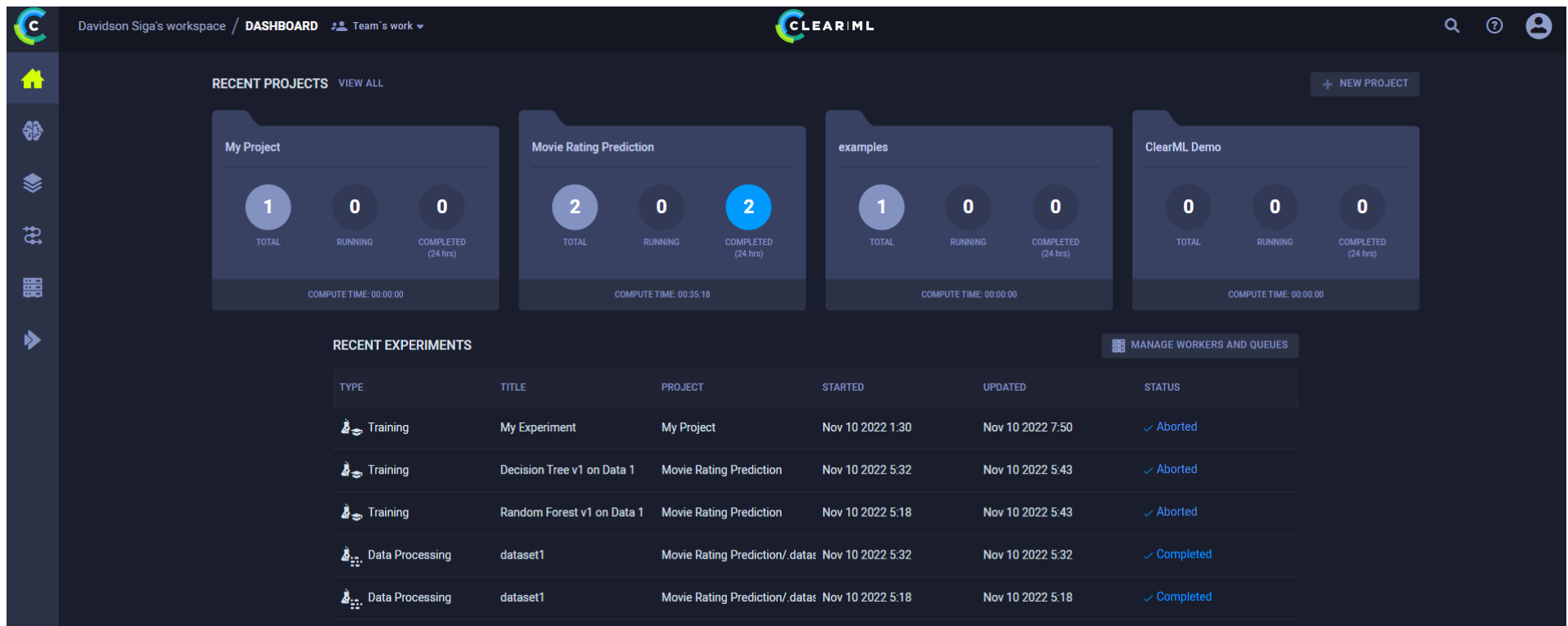
ML-Ops for efficient productionizing

Enter, Clear ML ☺ for ML Ops

Mission : Accelerate ML Adoption, making ML seamless to integrate into any Software and Hardware product out there



Easy to navigate UI





ClearML Integration

ClearML + Movie Rating Prediction

GitHub Walkthrough

https://github.com/sigadavid96/MLOPS_ClearML_MovieRatingPrediction

The screenshot shows a GitHub repository page for 'sigadavid96 / MLOPS_ClearML_MovieRatingPrediction'. The repository is marked as 'Private'. At the top, there are navigation links: '<> Code', 'Issues', 'Pull requests', 'Actions', 'Projects', 'Security', 'Insights', and 'Settings'. Below these, there are buttons for 'main' (selected), '1 branch', and '0 tags', along with 'Go to file', 'Add file', and a green 'Code' button. The repository description is 'Using the ClearML tools for ML Ops'. It shows 0 stars, 1 watching, and 0 forks. The file list includes 'Codes' (Create model_training.ipynb, 3 hours ago) and 'README.md' (Update README.md, 3 hours ago). The README content is visible, starting with 'Using Clear ML for ML Operations on a Movie Rating Prediction Problem'. It recommends creating a Conda Python Environment and provides instructions for setting up the environment in VS Code. The code snippet shown is: `conda create -n clearml python=3.9 anaconda`. On the right sidebar, there are sections for 'About', 'Releases' (No releases published, Create a new release), 'Packages' (No packages published, Publish your first package), and 'Languages' (Jupyter Notebook 100.0%).



ClearML

The Strengths and Limitations

Strengths and Weakness

Strengths :

- Very easy to integrate it to existing codes/system
- Very appealing & Easy to Use UI
- Very active & responsive community on Slack
- 100% Open Source Codes + Docker Hub

Weakness :

- Only supports Python and not other statistical languages like R
- The integration with SkLearn is only with JobLib and not as deep as Keras or Torch





Thank you!

Happy to connect :)

Davidson Siga
dsiga@cmu.edu

[linkedin.com/in/DavidsonSiga](https://www.linkedin.com/in/DavidsonSiga)
MISM 16, CMU

Carnegie Mellon University