

# The human loop in machine learning

Sayak Paul, Deep Learning Associate at [PyImageSearch](#)

DevFest Bhubaneswar 2019, September 22



2019



# Agenda

- The era of [AutoML](#): Automating the automation
- Machine learning: From a software engineer's keystrokes
- Demystifying the human involvement in machine learning projects
- A primer to [active learning](#)
- [Machine Learning-Assisted Humans](#) vs [Human-Assisted Machine Learning](#)
- Guiding lights

# Agenda

- The era of [AutoML](#): Automating the automation
- Machine learning: From a software engineer's keystrokes
- Demystifying the human involvement in machine learning projects
- A primer to [active learning](#)
- [Machine Learning-Assisted Humans](#) vs [Human-Assisted Machine Learning](#)
- Guiding lights
- No NeurIPS papers today!

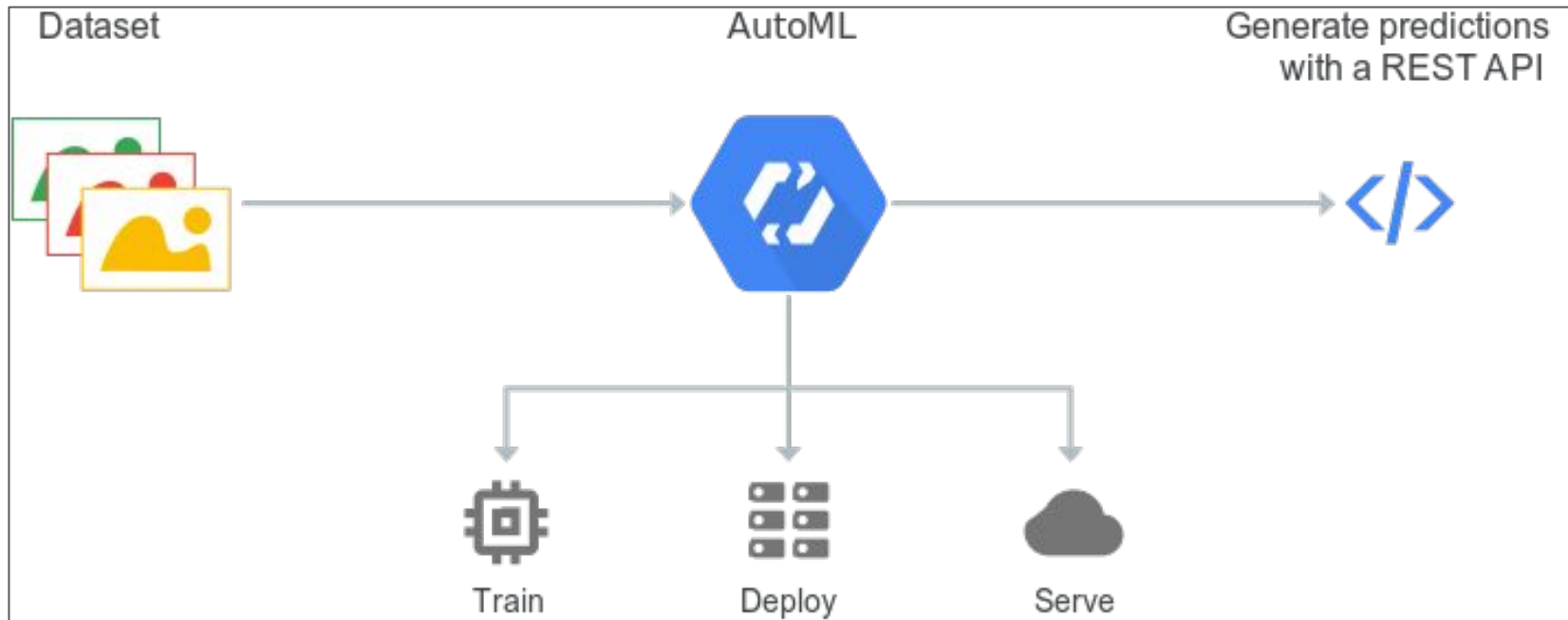
# A for AutoML

- “The automation of automating automation” - [Sebastian Raschka](#)
  - Automating the hyperparameter tuning process

# A for AutoML

- “The automation of automating automation” - [Sebastian Raschka](#)
  - Automating the hyperparameter tuning process
  - Given a dataset find the most suitable machine learning architecture ([NAS](#))

# A for AutoML



Source: <https://cloud.google.com/automl/>



Where do the humans  
stand then?

# Life-cycle of a machine learning project

- Problem [understanding](#)
- Data collection, wrangling and so on
- [Understanding](#) of the data
- Beginning the modeling process
- Evaluate, tune, repeat
- Model in production

**Note** that the above steps are *typical*



# Life-cycle of a machine learning project

- Problem [understanding](#)
- Data collection, wrangling and so on
- [Understanding](#) of the data
- Beginning the modeling process
- Evaluate, tune, repeat
- Model in production



Tableau, Facet, TensorBoard Projector ...

**Note** that the above steps are *typical*

# Life-cycle of a machine learning project

- Problem [understanding](#)
- Data collection, wrangling and so on
- [Understanding](#) of the data
- Beginning the modeling process
- Evaluate, tune, repeat
- Model in production

Tableau, Facet, TensorBoard Projector ...

Keras-tuner, HPARAMS ...

**Note** that the above steps are *typical*



**But how do we know if a problem is suitable for machine learning?**

Demystifying the human  
involvement in machine learning  
projects

# Problem framing is a scientific process

1. Frame the problem: What will traffic be like tomorrow?
2. Make a hypothesis: Weather forecast could be informative.
3. Collect the data: Collect historical traffic and weather data.
4. Test hypothesis: Test a model with the data
5. Analyze results: Is this model better than existing systems?
6. Reach a conclusion: I should (not) use this model, because of X, Y, and Z.
7. Refine and repeat: Time of year could be a helpful signal.



Can a machine automate this process?

# Using heuristics

- Is it possible to solve the problem without machine learning?

# Using heuristics

- Is it possible to solve the problem without machine learning?
- Why deploy a model if it cannot beat a heuristic?



# Using heuristics

- Is it possible to solve the problem without machine learning?
- Why deploy a model if it cannot beat a heuristic?
- Non-ML solutions can sometimes be *simpler* to maintain than ML solutions.

# An intro to Active Learning

# An intro to **Active Learning**

- The process of selecting the right data for human review:
  - Model predictions are no better than random guesses  
**(Uncertainty Sampling)**

# An intro to **Active Learning**

- The process of selecting the right data for human review:
  - Model predictions are no better than random guesses  
(**Uncertainty Sampling**)
  - Unlabeled data points that are in very different parts of the problem space (**Diversity Sampling**)

# An intro to **Active Learning**

- The process of selecting the right data for human review:
  - Model predictions are no better than random guesses  
(**Uncertainty Sampling**)
  - Unlabeled data points that are in very different parts of the problem space (**Diversity Sampling**)

It helps improving the overall **speed** and **cost** of training data!

Machine Learning-Assisted

Humans **vs** Human-Assisted

Machine Learning

# Machine learning & human involvement

- Making a Machine Learning application more accurate with human input

# Machine learning & human involvement

- Making a Machine Learning application more accurate with human input
- Improving a human task with the aid of Machine Learning



# Wrapping up

- Automated machine learning
- More human involvement in machine learning
- The sweet rivalry

# See you next time



Find me here:

[sayak.dev](https://sayak.dev)

Thank you very much :)



Experts

