



SageMaker Experiments

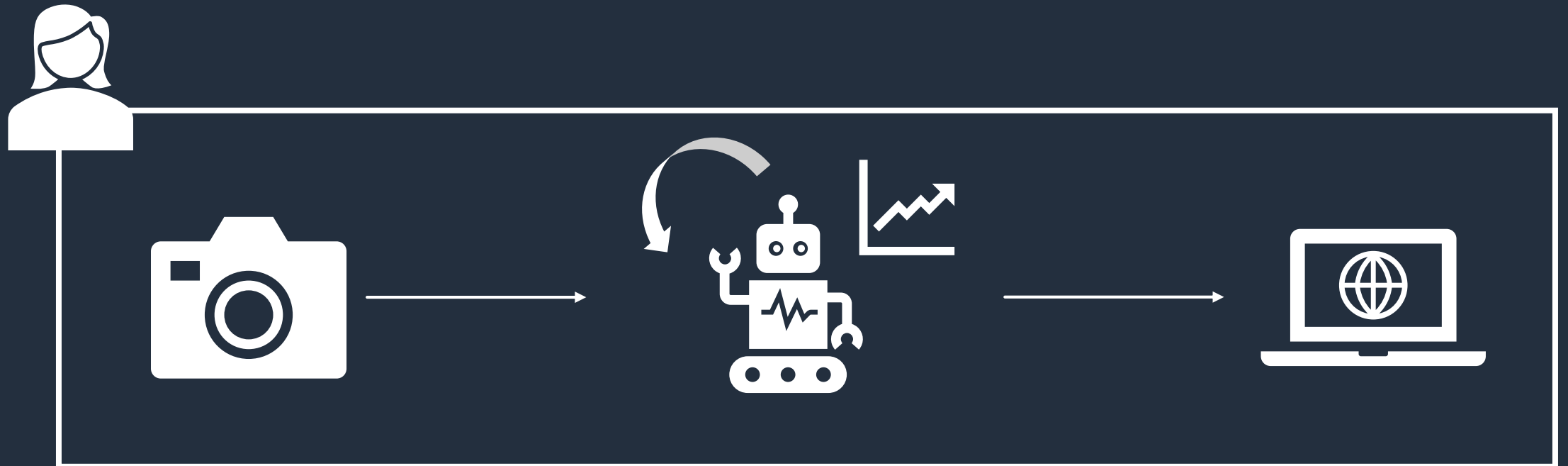
Create, manage, analyze, compare and track experiments at scale

Pranav Murthy

Sr. AI/ML Specialist SA

What is Machine Learning Experimentation?

Machine Learning Experimentation is an iterative process of trying different approaches to a machine learning problem in order to find the best one. This involves testing different algorithms, hyperparameters, and data sets to see which combination gives the best results.



Amazon SageMaker Experiments

Create, manage, analyze, compare and track experiments at scale



Tracking at scale

Track parameters and metrics across experiments and users



Custom organization

Organize experiments by teams, goals, and hypotheses



Visualization

Easily visualize experiments and compare



Metrics and logging

Log custom metrics using the Python SDK and APIs



Fast iteration

Quickly go back and forth, and maintain high-quality

Amazon SageMaker Experiments

Create, manage, analyze, compare and track experiments at scale



Amazon SageMaker Experiments

Create, manage, analyze, compare and track experiments at scale

Key Concepts

- 1. Experiment:** *An experiment is a collection of runs. When you initialize a run in your training loop, you include the name of the experiment that the run belongs to. Experiment names must be unique within your AWS account*
- 2. Run:** *A run consists of all the inputs, parameters, configurations, and results for one iteration of model training. Initialize an experiment run for tracking a training job with `Run()`*

Amazon SageMaker Experiments

Create, manage, analyze, compare and track experiments at scale

Integration with SageMaker Clarify

```
with Run(
    experiment_name='clarify-experiment',
    run_name="joint-run",
    sagemaker_session=sagemaker_session,
) as run:
    xgb.fit({"train": train_input}, logs=False)
    clarify_processor.run_bias(
        data_config=bias_data_config,
        bias_config=bias_config,
        model_config=model_config,
        model_predicted_label_config=predictions_config,
        pre_training_methods="all",
        post_training_methods="all",
    )
    clarify_processor.run_explainability(
        data_config=explainability_data_config,
        model_config=model_config,
        explainability_config=shap_config,
    )
```

Amazon SageMaker Experiments

Create, manage, analyze, compare and track experiments at scale

Integration with SageMaker Clarify

```
with Run(  
    experiment_name='clarify-experiment',  
    run_name="joint-run",  
    sagemaker_session=sagemaker_session,  
) as run:  
    xgb.fit({"train": train_input}, logs=  
clarify_processor.run_bias(  
    data_config=bias_data_config,  
    bias_config=bias_config,  
    model_config=model_config,  
    model_predicted_label_config=pre  
pre_training_methods="all",  
post_training_methods="all",  
)  
clarify_processor.run_explainability  
data_config=explainability_data_  
model_config=model_config,  
explainability_config=shap_conf:  
)
```

Bias Report

clarify-experiment-joint-run

Reports

The computed bias metrics are below:
Predicted column: Target
Predicted value or threshold: 1
Column analyzed for bias: Sex

Bias metric	Bias value	Description
Conditional Demographic Disparity in Labels (CDDL)	0.21	The metric examines whether, in the training data, the disadvantaged class h...
Class Imbalance (CI)	0.35	Detects if the advantaged group is represented in the dataset at a substantia...
Difference in Positive Proportions in Labels (DPL)	0.2	Detects if one class has a significantly higher proportion of desirable (or, alte...
Jensen-Shannon Divergence (JS)	0.031	JS measures how much the label distributions of different classes diverge fro...
Kullback-Liebler Divergence (KL)	0.14	In a binary case, a relative entropy measure of how much the label distributi...
Kolmogorov-Smirnov Distance (KS)	0.2	This metric is equal to the maximum divergence in a label across the classes ...
L-p Norm (Lp)	0.28	This measure of distance in label distributions is the normed direct distance ...
Total Variation Distance (TVD)	0.2	This measure of distance in label distributions is half the Hamming distance ...
Accuracy Difference (AD)	-0.11	This metric examines whether the classification by the model is more accurat...
Conditional Demographic Disparity in Predicted Labels (CDDPL)	0.2	The metric examines whether the model predicted a bigger proportion of rej...
Difference in Acceptance Rates (DAR)	-0.0074	The difference in the rates of positive predicted outcomes across the advant...
Difference in Conditional Acceptance (DCA)	-0.23	This metric compares the actual labels to the predicted labels from the mod...
Difference in Conditional Outcomes (DCR)	0.13	This metric compares the actual labels to the predicted labels from the mod...
Disparate (Adverse) Impact (DI)	0.33	This metric examines whether the model predicts outcomes differently for e...
Difference in Positive Proportions in Predicted Labels (DPPL)	0.092	This metric examines whether the model predicts outcomes differently for e...

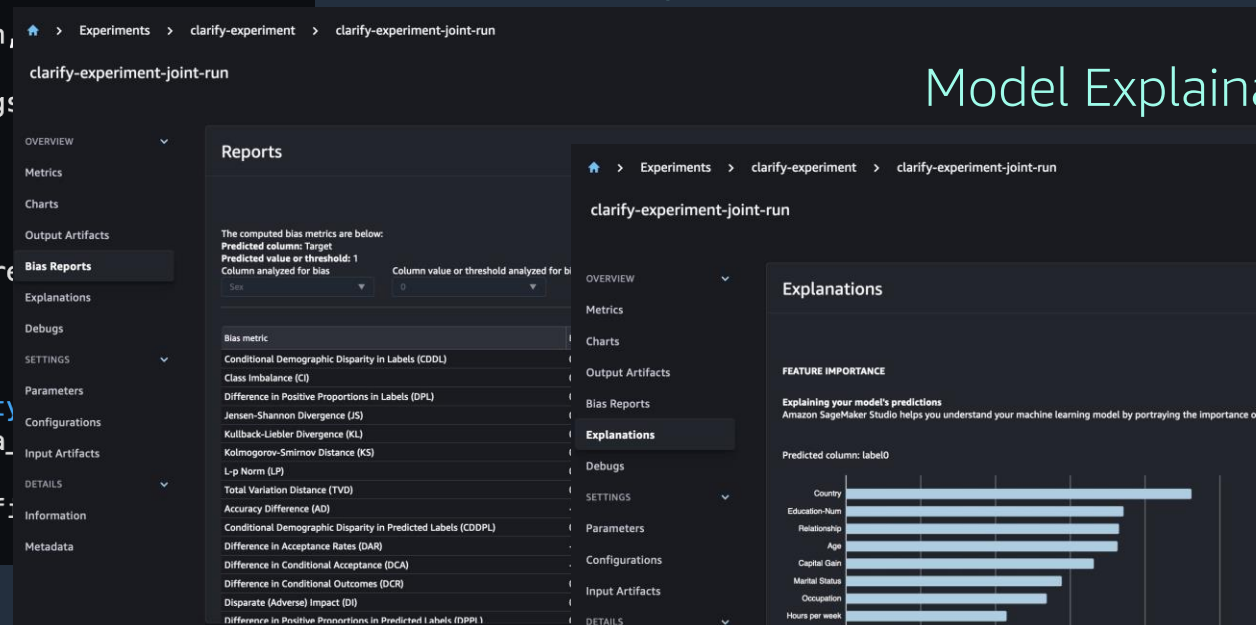
Amazon SageMaker Experiments

Create, manage, analyze, compare and track experiments at scale

Integration with SageMaker Clarify

```
with Run(  
    experiment_name='clarify-experiment',  
    run_name="joint-run",  
    sagemaker_session=sagemaker_session,  
) as run:  
    xgb.fit({"train": train_input}, logs=  
clarify_processor.run_bias(  
    data_config=bias_data_config,  
    bias_config=bias_config,  
    model_config=model_config,  
    model_predicted_label_config=pre  
pre_training_methods="all",  
post_training_methods="all",  
)  
clarify_processor.run_explainability  
data_config=explainability_data_  
model_config=model_config,  
explainability_config=shap_conf:  
)
```

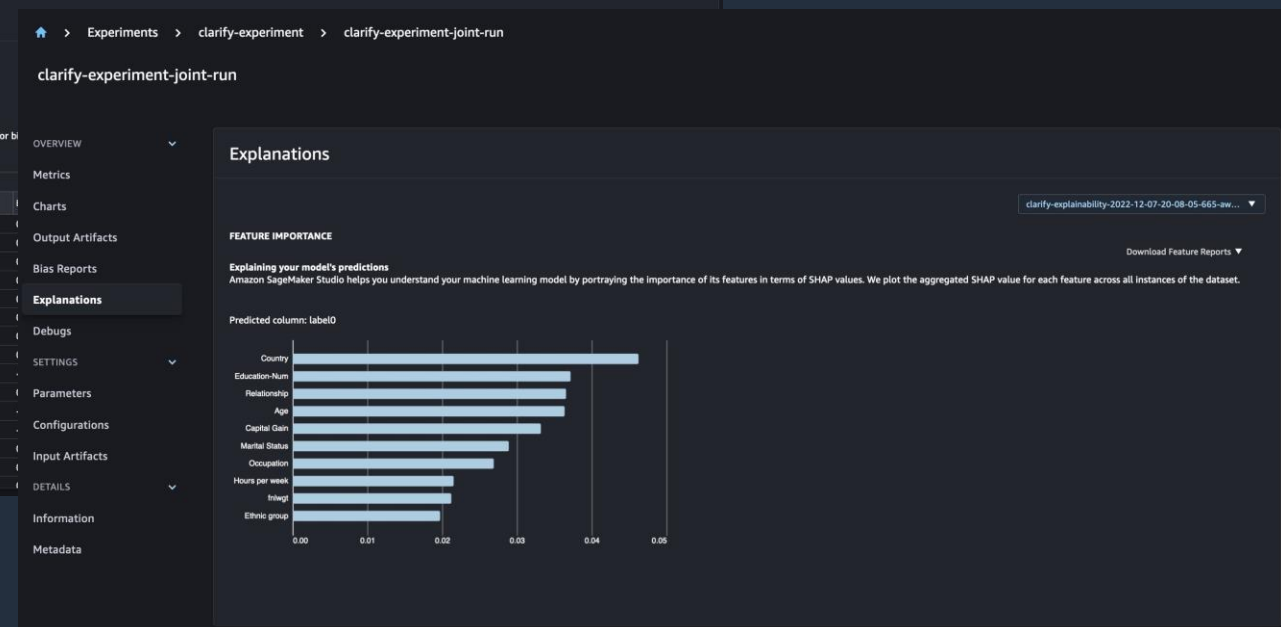
Bias Report



The screenshot shows the SageMaker Clarify Bias Report interface. The breadcrumb navigation is "Experiments > clarify-experiment > clarify-experiment-joint-run". The left sidebar has "Bias Reports" selected. The main content area shows "Reports" with a table of bias metrics. The table has columns for "Bias metric", "Predicted column: Target", "Predicted value or threshold: 1", and "Column value or threshold analyzed for bias". The table lists various metrics such as Conditional Demographic Disparity in Labels (CDDL), Class Imbalance (CI), and Accuracy Difference (AD).

Bias metric	Predicted column: Target	Predicted value or threshold: 1	Column value or threshold analyzed for bias
Conditional Demographic Disparity in Labels (CDDL)	Sex	0	
Class Imbalance (CI)			
Difference in Positive Proportions in Labels (DPL)			
Jensen-Shannon Divergence (JS)			
Kullback-Liebler Divergence (KL)			
Kolmogorov-Smirnov Distance (KS)			
L-p Norm (Lp)			
Total Variation Distance (TVD)			
Accuracy Difference (AD)			
Conditional Demographic Disparity in Predicted Labels (CDDPL)			
Difference in Acceptance Rates (DAR)			
Difference in Conditional Acceptance (DCA)			
Difference in Conditional Outcomes (DCR)			
Disparate (Adverse) Impact (DI)			
Difference in Positive Proportions in Predicted Labels (DPPL)			

Model Explainability



The screenshot shows the SageMaker Clarify Model Explainability interface. The breadcrumb navigation is "Experiments > clarify-experiment > clarify-experiment-joint-run". The left sidebar has "Explanations" selected. The main content area shows "Explanations" with a "FEATURE IMPORTANCE" section. It includes a horizontal bar chart showing the aggregated SHAP values for various features. The x-axis ranges from 0.00 to 0.06. The y-axis lists features: Country, Education-Num, Relationship, Age, Capital Gain, Marital Status, Occupation, Hours per week, Inwtgt, and Ethnic group. Country has the highest importance, followed by Education-Num and Relationship.

Feature	Importance (SHAP value)
Country	~0.055
Education-Num	~0.045
Relationship	~0.042
Age	~0.038
Capital Gain	~0.035
Marital Status	~0.032
Occupation	~0.028
Hours per week	~0.025
Inwtgt	~0.022
Ethnic group	~0.018





Thank you!

Pranav Murthy

<https://linkedin.com/in/pranav-murthy-6bbb5773/>