AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

AIM322

Gain greater visibility and transparency into your ML models

Chris Knoch (he/him)
Principal Product Manager
AWS

Pinar Yilmaz (she/her)
Principal Engineer
AWS

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Responsible AI overview

Amazon SageMaker Clarify life cycle of responsible AI

Deep dive demo

New features overview

Q&A

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The need for responsible AI

- Ethical challenges posed by AI systems
- Inherent biases present in society
- Reflected in training data
- AI/ML models prone to amplifying such biases

Weapons of Math Destruction

How Big Data Increases Inequality and Threatens Democracy

Cathy O'Neil

The Ethical Algorithm: The Science of Socially Aware Algorithm Design

Michael Kearns and Aaron Roth

The Black Box Society

The Secret Algorithms that Control Money and Information

Frank Pasquale

Algorithms of Oppression

How Search Engines Reinforce Racism

Safiya Umoja Noble

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Laws against discrimination

Citizenship

Immigration Reform and Control Act



Sex

Equal Pay Act of 1963;
Civil Rights Act of 1964

Age

Age Discrimination in Employment Act of 1967

Race

Civil Rights Act of 1964

Disability status

Rehabilitation Act of 1973;
Americans with Disabilities Act of 1990

And more...

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Fairness

Privacy



Transparency



Explainability

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Responsible AI motivation & opportunities

Regulatory: Certain regulators want to understand why the ML model made a given prediction, and also whether the predictions made were free from bias in training and at inference

Business: Providing explanations to internal teams (loan officers, customer service rep, compliance teams) and end users/customers

Data science: Improving models through better feature engineering and training data generation, understanding failure modes of the model, debugging model predictions, etc.

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Responsible AI challenges

Explainability is critical for responsible AI:

- Machine learning models are often complex & opaque
- Regulations may require companies to be able to explain model predictions
- Internal stakeholders and customers may need explanations for model behavior
- Data science teams can improve models if they understand model behavior

Bias detection is necessary at multiple phases:

- Imbalances and bias may be present in the initial training dataset
- Bias may develop during training
- Bias may develop over time after model deployment

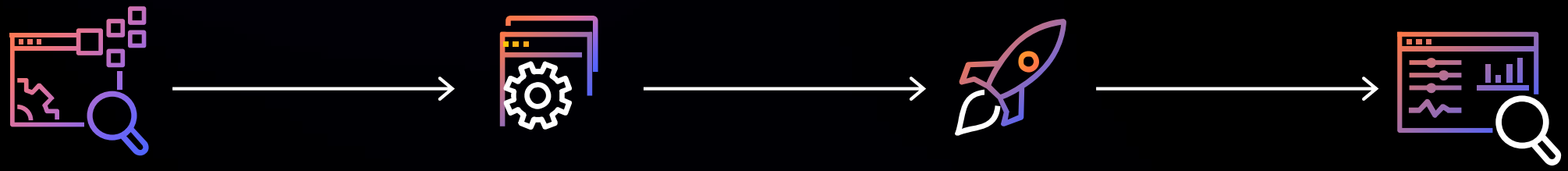
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Lessons learned

- Fairness as a process
  - Notions of bias & fairness are highly application dependent
  - Choice of the attribute(s) for which bias is to be measured & the choice of the bias metrics to be guided by social, legal, and other non-technical considerations
- Collaboration/consensus across key stakeholders
- Wide spectrum of customers with different levels of technical background
  - Managed service versus open-source packages
- Monitoring of the deployed model
- Fairness & explainability considerations across the ML life cycle

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Responsible AI in the ML life cycle



Build

Perform bias analysis during exploratory data analysis

Train

Conduct bias and explainability analysis after training

Deploy

Explain individual inferences for model in production

Monitor

Validate bias and relative feature importance over time

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Clarify: Responsible AI & explainability service

SageMaker Clarify has two key feature sets threaded through the entire SageMaker ML life cycle:

Explainability for model understanding:

Understand which features are most influential on the overall model

Explain individual predictions (Example: “Why didn’t the model approve a loan?”)

Business Action: XAI enables non-data scientist to understand and act on data

Model monitor integration to detect drift in feature important over time

Bias detection for responsible AI

Bias detection: Help customers understand whether their data (before training) or their trained model have bias across different attributes that a customer can specify (i.e gender, zip code, income, etc.)

Model monitor integration to detect drift over time in the bias of the model predictions

Regulatory environment: Helps customers satisfy the requirements of their internal risk and external regulatory bodies by providing exportable reports for all of the metrics we generate

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Clarify

Explain machine learning models and assess your data and models for bias



Detect bias during data preparation

Detect bias across attributes in your training data



Detect bias in your trained model

Evaluate the degree to which various types of bias are present in your model



Explain overall model behavior

Understand the relative importance of each feature to your model's behavior



Explain individual predictions

Understand the relative importance of each feature for individual inferences



Detect drift in bias and model behavior over time

Provide alerts and detect drift over time due to changing real-world conditions



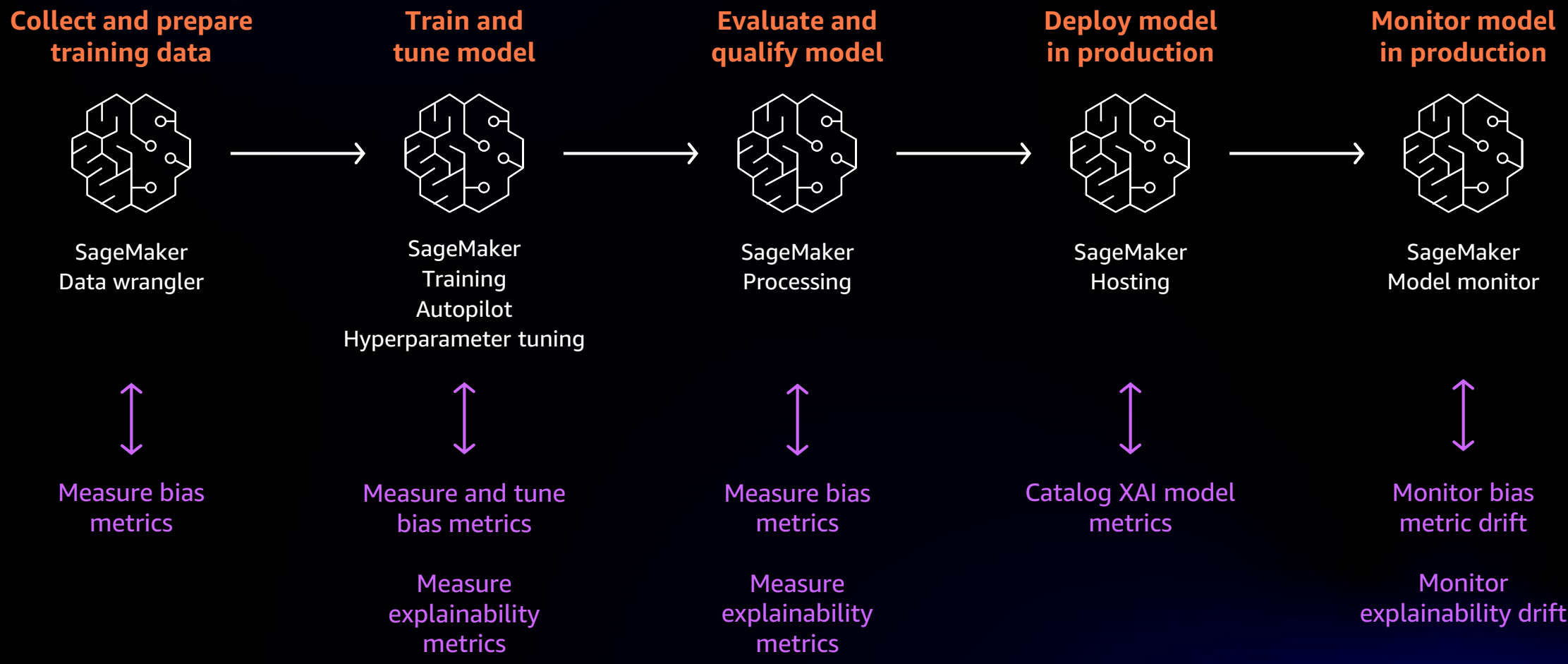
Provide holistic reports to your auditors

Offer internal risk teams and external regulatory bodies the visibility they need

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Clarify works across the ML life cycle on multiple SageMaker services



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer spotlight

Zopa

Zopa is a digital bank that uses machine learning to identify possible fraud in their loan applications. Suspicious loan requests need a manual underwriter review (current SLA = 1 day). They need explainability for their internal compliance and data science teams and for future regulatory requirements.

> In our machine learning applications, such as fraud detection, it is important for us to understand how each factor contributes to the model's decision. Having visibility into the reasoning of the model gives confidence to our stakeholders, both internal and external. It also helps our operations team respond faster and provide better service to our customers. With Amazon SageMaker Clarify, we can now produce model explanations more quickly and seamlessly."
>
> **Jiahang Zhong, Head of Data Science, Zopa**

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer spotlight

Varo

Varo is an online-only bank, founded to target millennials and younger people who are disillusioned with traditional banks. In July 2020 this customer became the first fintech company to obtain a national bank charter. To operate independently, the customer was required to establish a full compliance regime. The customer is using Amazon SageMaker Clarify's global (model averages) and local (per-prediction) explanations for risk segmentation with their xgboost model. They are planning to use Amazon SageMaker Clarify's explainability techniques to explain results of machine learning models for additional use cases like underwriting, limits, and personalization.

> "Varo has a strong commitment to the explainability and transparency of our ML models, and we're excited to see the strong results from Amazon SageMaker in advancing these efforts."
>
> **Sachin Setty, Head of Data Science, Varo Bank**

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer spotlight



Bundesliga

Since naming AWS as its official technology provider in January 2020, the DFL Deutsche Fußball Liga – organizer and marketer of Germany's top soccer leagues Bundesliga and Bundesliga 2 – and AWS have embarked on a journey together to bring advanced sports analytics to life for fans and TV broadcasters around the globe. Bundesliga Match Facts powered by AWS provide a more engaging fan experience during soccer matches for Bundesliga fans around the world. One such example is the xGoals (short for "Expected Goals") statistic which uses an ML model trained on Amazon SageMaker. It is used to determine real-time goal-scoring chances, based on 16 different factors in any given situation during a match, such as the distance and angle of the striker to the goal, current player running speed, the number of surrounding opponents, the distance of the keeper from the goal line, amongst others.

> *"**Amazon SageMaker Clarify seamlessly integrates with the rest of the Bundesliga Match Facts digital platform and is a key part of our long-term strategy of standardizing our ML workflows on Amazon SageMaker**. By using AWS's innovative technologies such as machine learning to deliver more in-depth insights and provide fans a better understanding of the split-second decisions made on the pitch, Bundesliga Match Facts enables viewers to gain deeper insights into the key decisions in each match."*
>
> Andreas Heyden, Executive Vice President of Digital Innovations for the DFL Group

aws

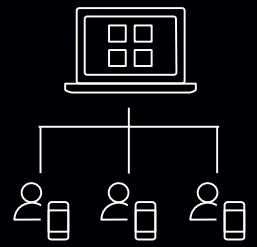
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Popular use cases for explainability (XAI) & bias



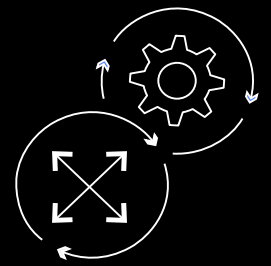
Regulatory requirements

Understand why ML models made a given decision and whether the decision was free from bias, both in training and at inference



Internal reports & customer needs

Provide explanations to internal teams (loan officers, customer service reps, forecasting teams) and end users or customers

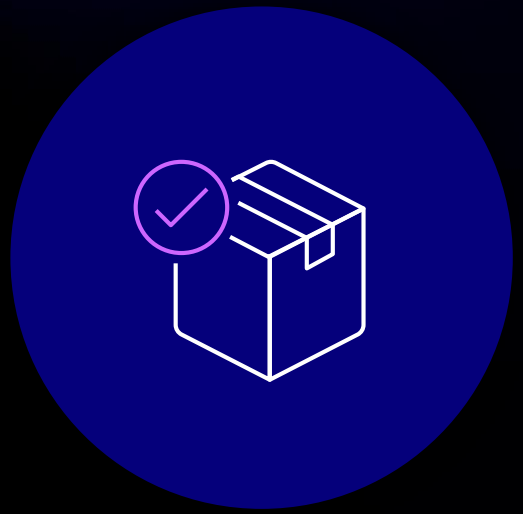


ML model optimization

Improve model quality with a deeper understanding of whether a model is making the right inferences for the right reasons

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

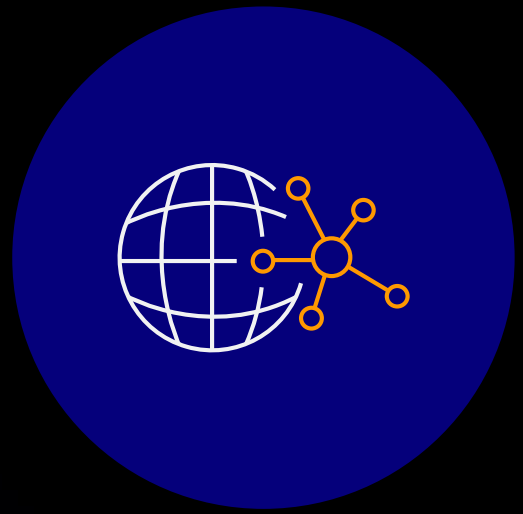
SageMaker Clarify – Pricing and availability



SageMaker Clarify is generally available today



SageMaker Clarify is available at no additional cost as part of Amazon SageMaker



SageMaker Clarify is available in all AWS Regions where SageMaker is available

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.