😊 Amazon Fraud Detector

Amazon Fraud Detector is a fully managed machine learning service by AWS that helps businesses identify and prevent online fraud in real-time. It simplifies the process of building and deploying machine learning models to detect suspicious activities like online payment fraud, fake account creation, and identity theft without requiring machine learning expertise.

Key Features:

1. **Fraud Detection Models**:
    - Amazon Fraud Detector provides pre-built fraud detection models trained on Amazon's own data from years of combating fraud on its e-commerce platform. These models are highly accurate and can detect fraudulent activities with minimal configuration.
2. **Custom Model Training**:
    - Users can upload their own historical data to train custom fraud detection models. The service will automatically handle the complexities of model development, including feature engineering, model selection, and tuning.
3. **Real-Time Predictions**:
    - Once deployed, Amazon Fraud Detector evaluates events like account creation or transaction attempts and returns fraud risk predictions in real time, allowing businesses to take immediate actions (e.g., flagging suspicious accounts or blocking transactions).
4. **Fraud Risk Scores and Outcomes**:
    - The service provides a fraud risk score for each event and allows users to define outcomes based on these scores. For instance, if the score is above a certain threshold, the system might trigger a manual review, send an alert, or block the transaction.
5. **Customizable Rules Engine**:
    - Users can combine machine learning-based fraud predictions with custom business rules to better fine-tune fraud detection according to their specific use cases. For example, they might flag all transactions from certain locations as higher risk.
6. **Integration with AWS Services**:
    - Amazon Fraud Detector integrates seamlessly with other AWS services like Amazon S3 for data storage, AWS Lambda for triggering actions based on predictions, and Amazon CloudWatch for monitoring and logging.
7. **Fraud Detection Use Cases**:
    - Common fraud detection scenarios include:

- **Online payment fraud**: Detect suspicious transactions in e-commerce or financial services.
- **Account takeovers**: Identify unusual login behavior or attempts to access accounts using stolen credentials.
- **Fake account creation**: Prevent the creation of fake or fraudulent accounts used for spamming or other malicious activities.
- **Loyalty fraud**: Detect suspicious behavior in loyalty or rewards programs.

Benefits:

- **No Machine Learning Expertise Required**: Businesses can easily implement fraud detection solutions without having to build complex machine learning models from scratch.
- **Pre-Trained Models**: Amazon's models are built from years of experience and data, so businesses benefit from a sophisticated fraud detection system right out of the box.
- **Customizable to Your Business**: You can customize the system to suit specific business needs by uploading your own data or defining business-specific fraud detection rules.
- **Scalable**: Fraud Detector scales automatically to handle large volumes of transactions and events.
- **Real-Time Detection**: Allows for instant decision-making on potentially fraudulent activities.

How it Works:

1. **Data Collection**: The business provides historical event data (such as previous transactions, account registrations, or fraud outcomes) to Amazon Fraud Detector.
2. **Model Training**: The service either uses pre-built models or trains custom models on the provided data to learn patterns indicative of fraud.
3. **Deploy and Monitor**: Once deployed, the model can evaluate real-time events and produce fraud risk scores. These scores are used to trigger actions, such as blocking transactions or flagging them for review.

Integration:

- **Amazon S3**: Store and retrieve your event data for model training.
- **Amazon SNS**: Notify teams or systems when potential fraud is detected.
- **AWS Lambda**: Automate actions based on fraud detection, such as disabling user accounts or flagging suspicious transactions.
- **Amazon CloudWatch**: Monitor and log the performance of fraud detection models.

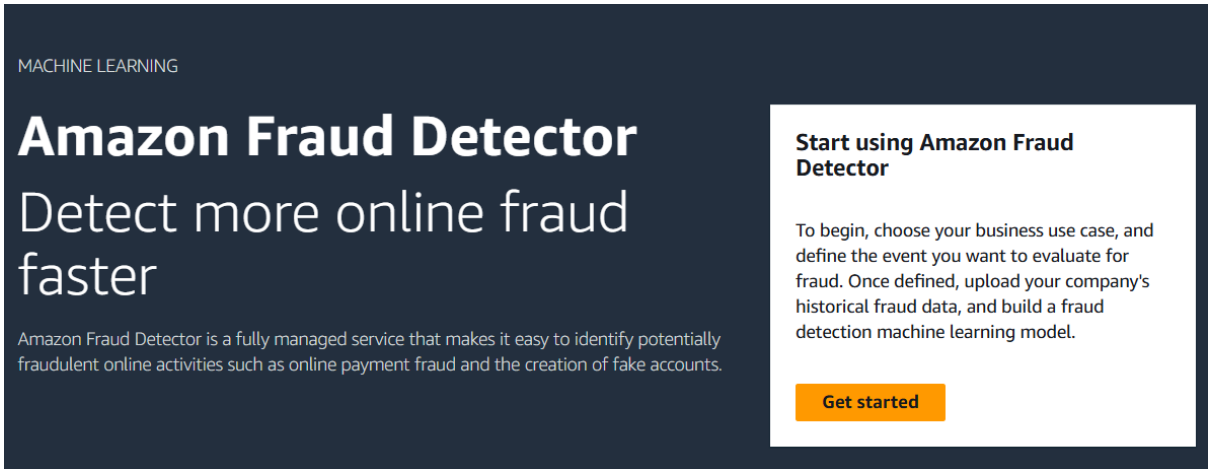
Common Use Cases:

- **Financial Services**: Detect fraudulent credit card transactions or suspicious account activity.
- **E-commerce**: Prevent payment fraud and protect customer accounts from hacking attempts.
- **Insurance**: Identify fraudulent claims or attempts to exploit insurance systems.
- **Gaming**: Detect fraudulent behavior in gaming platforms such as fake accounts or cheating.

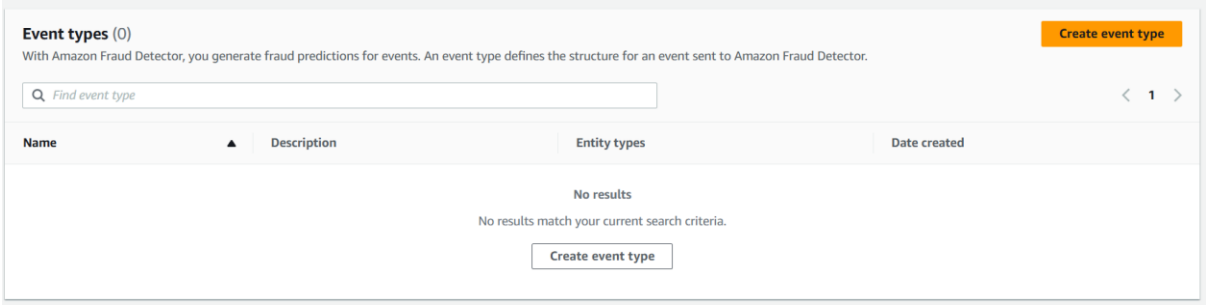
In summary, Amazon Fraud Detector provides a highly effective, scalable, and easy-to-use solution for detecting and mitigating fraud using machine learning. It helps businesses protect their operations and customers by reducing the likelihood of fraud losses.

😄 To begin with the Lab:

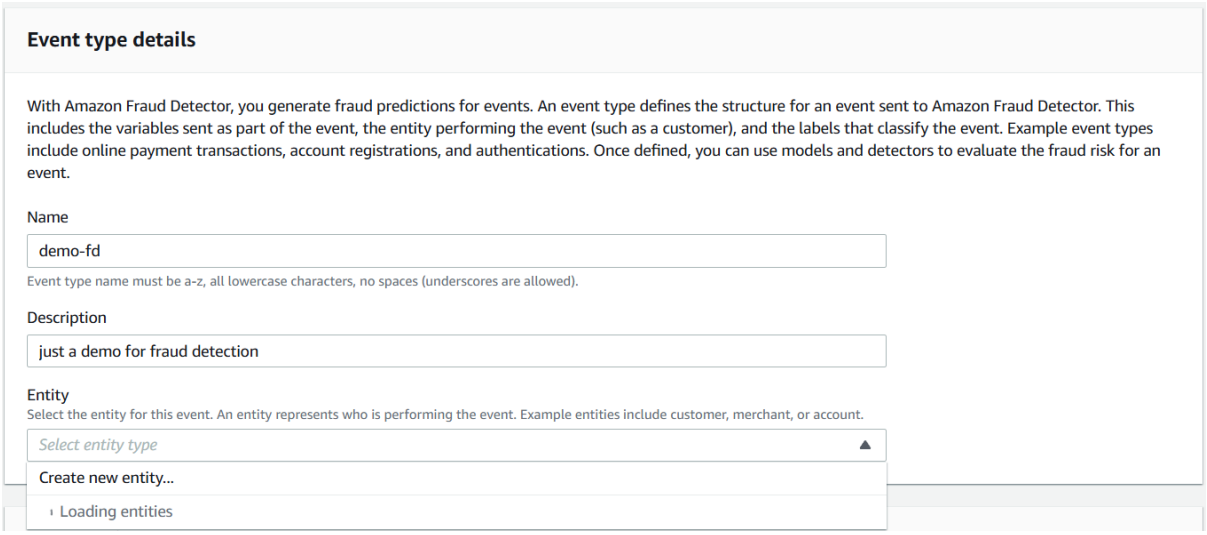
1. In your AWS Console, search for Amazon Fraud Detector and navigate to it. Click on Get started. Just remember that the region used in this service is Singapore.



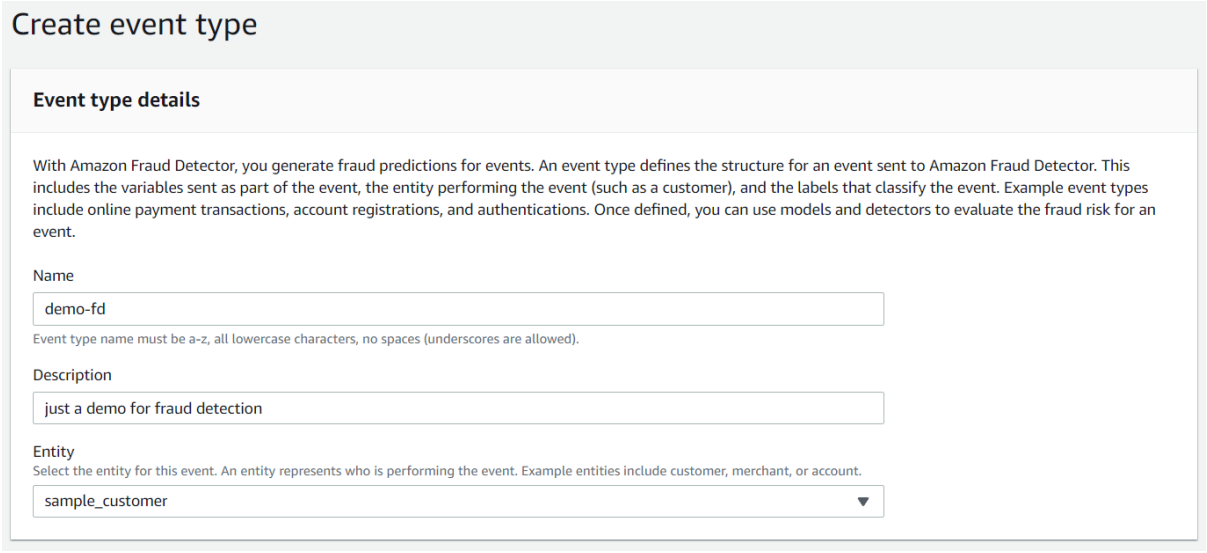
2. So, to begin with the fraud detector from the left pane choose events and click on creae event type.



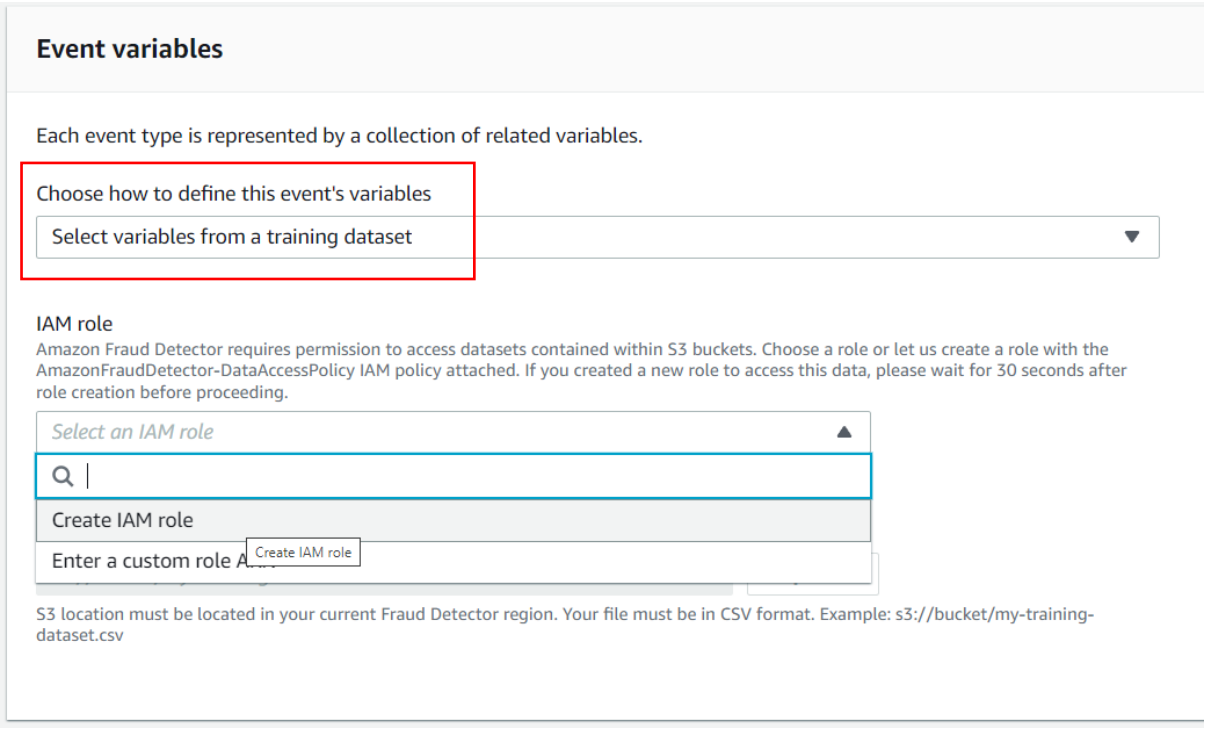
3. Here you need to give it a name and description then choose to create a new entity.



4. Below you can see that our entity has been created.

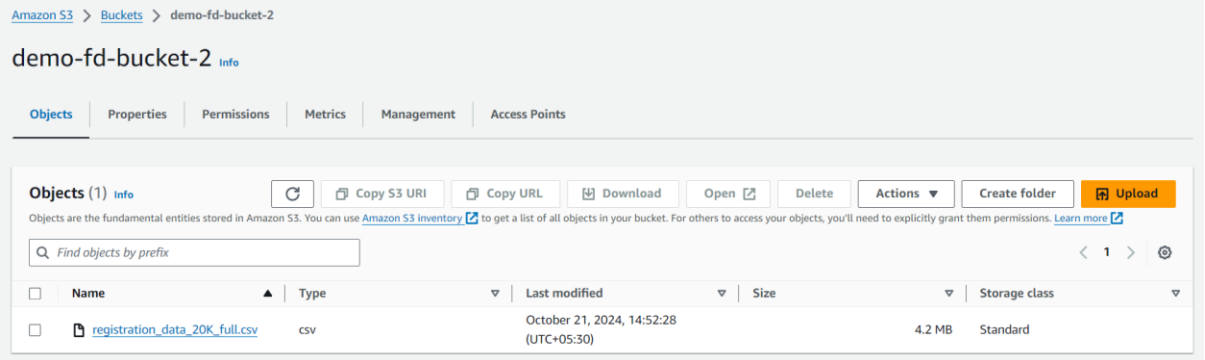


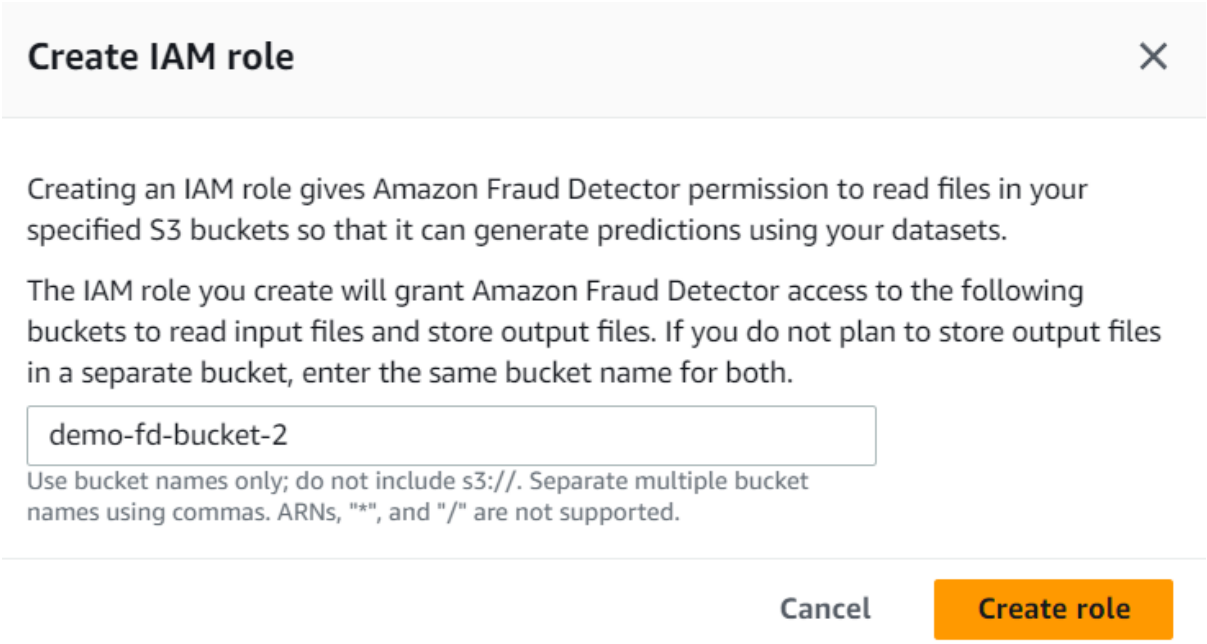
5. Then in the event variables section you need to choose the event highlighted below in the snapshot and then choose to create a new IAM role.



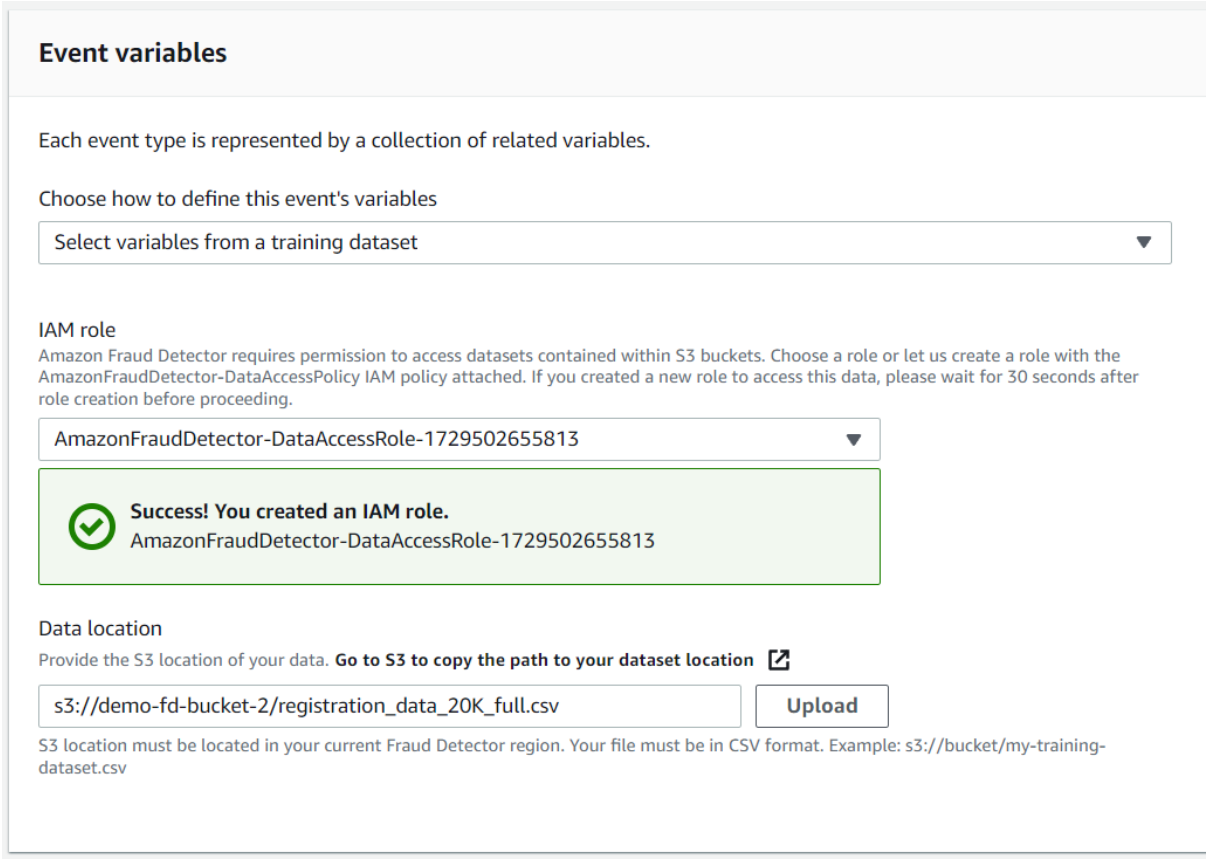
6. Here you can see that it is saying that you need to provide the name of your bucket.
7. Here you need to create an S3 bucket with a unique name and then upload the training data which you can get from the AWS GitHub repository using the link given down below.

   https://github.com/aws-samples/aws-fraud-detector-samples/blob/master/data/registration_data_20K_full.csv





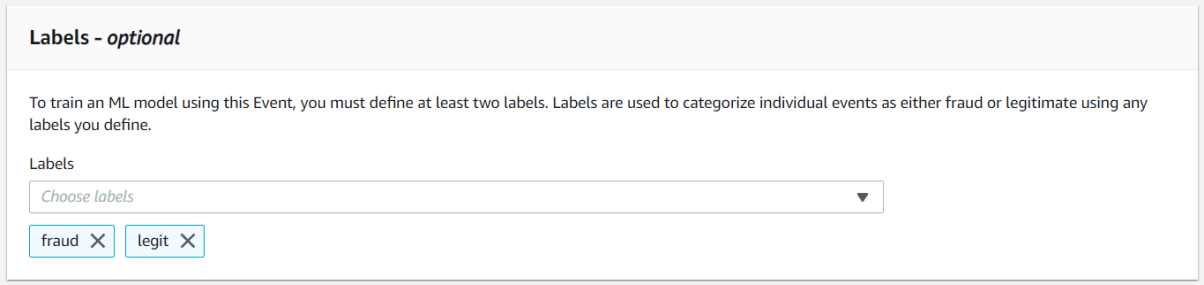
8. Now here you need to provide the Data location of your object so from your S3 bucket you need to copy the S3 object URI and paste it here then **click on Upload**.



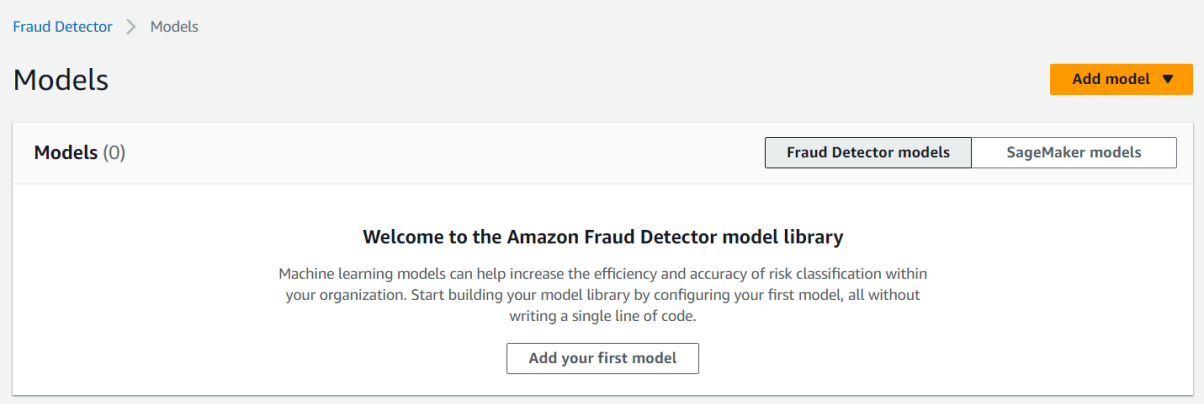
9. After click on Upload, you will see that you have a new section for Variable and variable types. So, fill the same variable types as you can see below.

Variable (7)	Variable type	
ip_address	IP Address	Remove
email_address	Email Address	Remove
billing_state	Billing Address: State or Province	Remove
user_agent	User Agent	Remove
billing_postal	Billing Address: Zip Code	Remove
phone_number	Phone Number	Remove
billing_address	Billing Address: Address Line 1	Remove

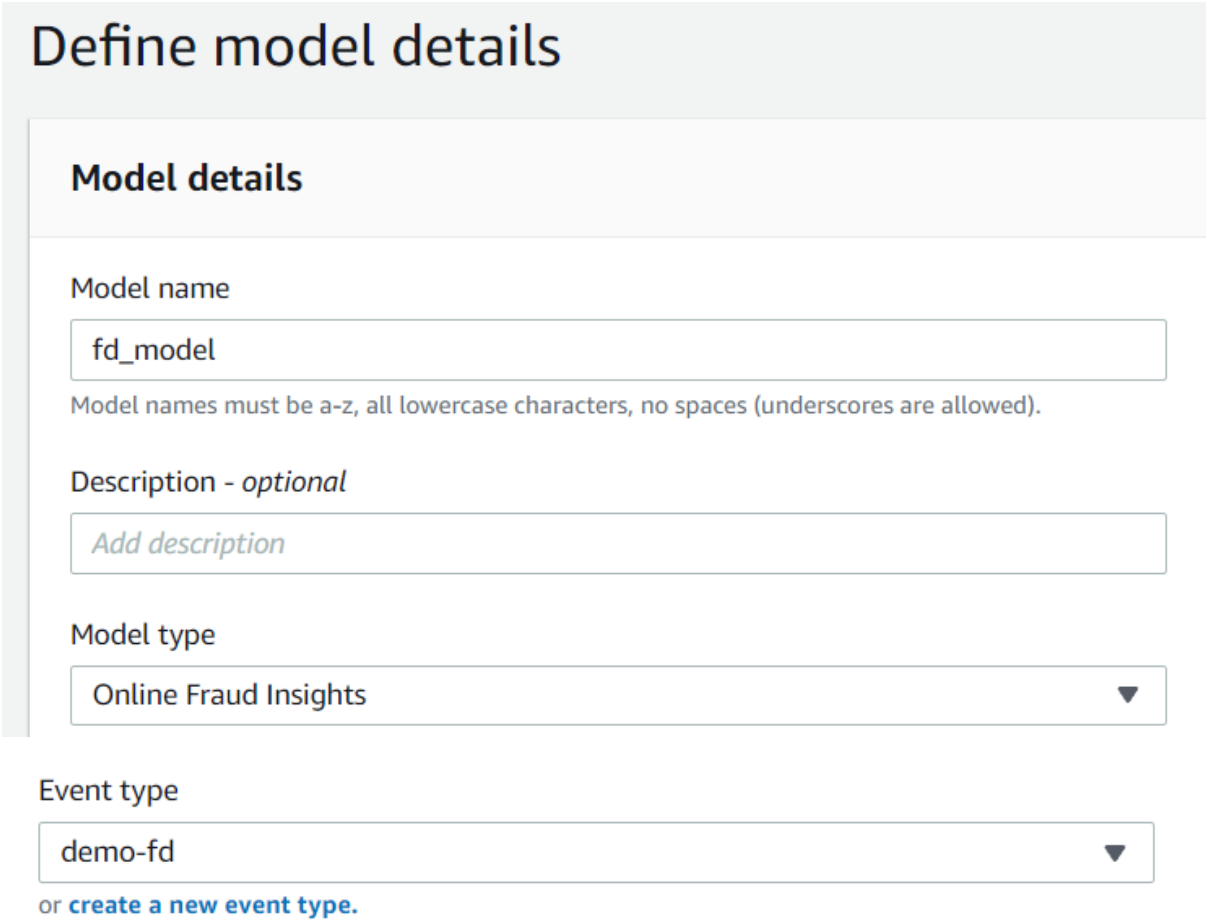
10. Then in the labels you need to create two labels for fraud and legit as you can see below.
11. After that move forward and click on create event.



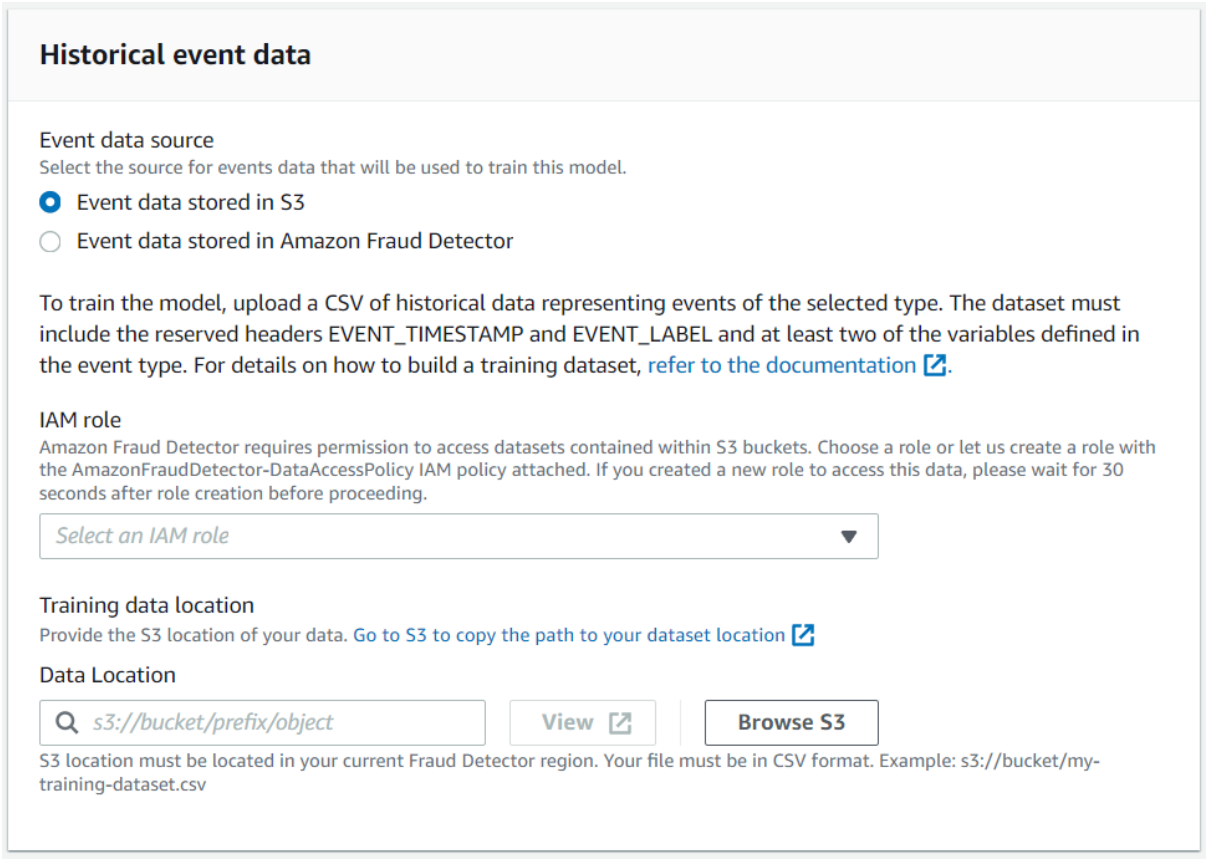
12. So, once your event has been created now you need to create a model for it.
13. From the left pane choose models and click on add model.



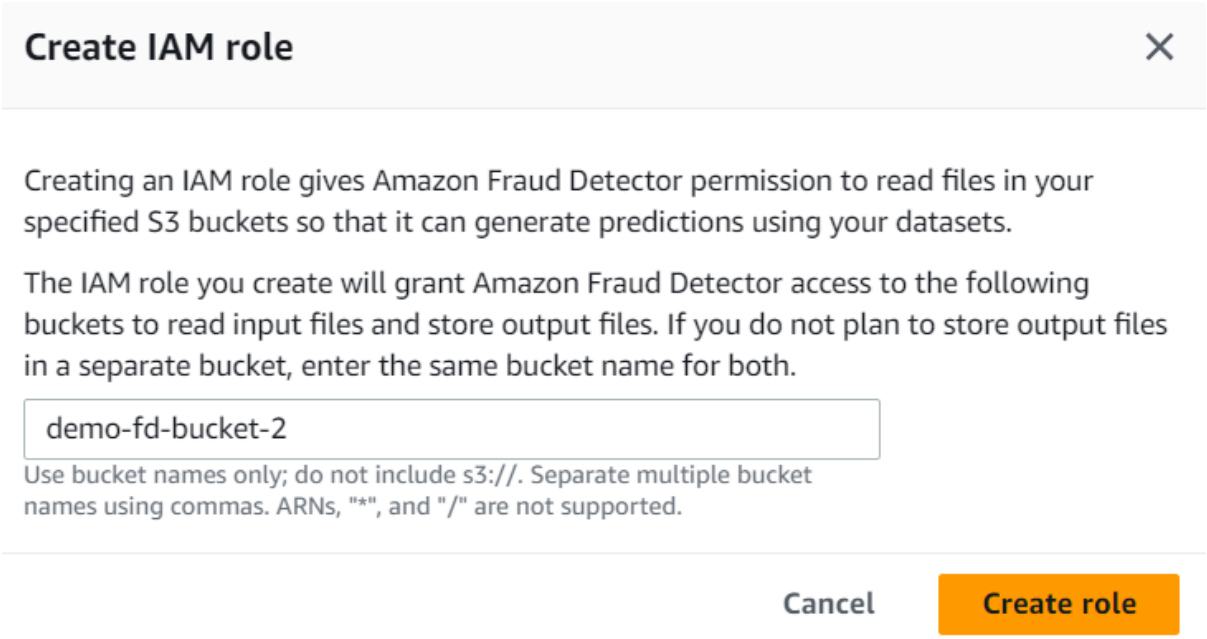
14. Here you need to give it a name and scroll down to event type and choose your recently created event.



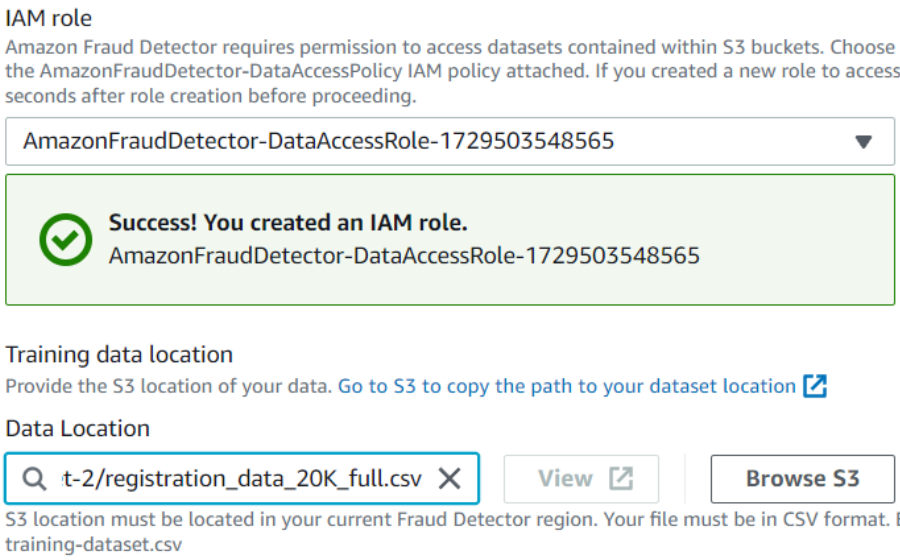
15. In the historical event data, you need to choose event data stored in S3. For the IAM role you need to create a new role.



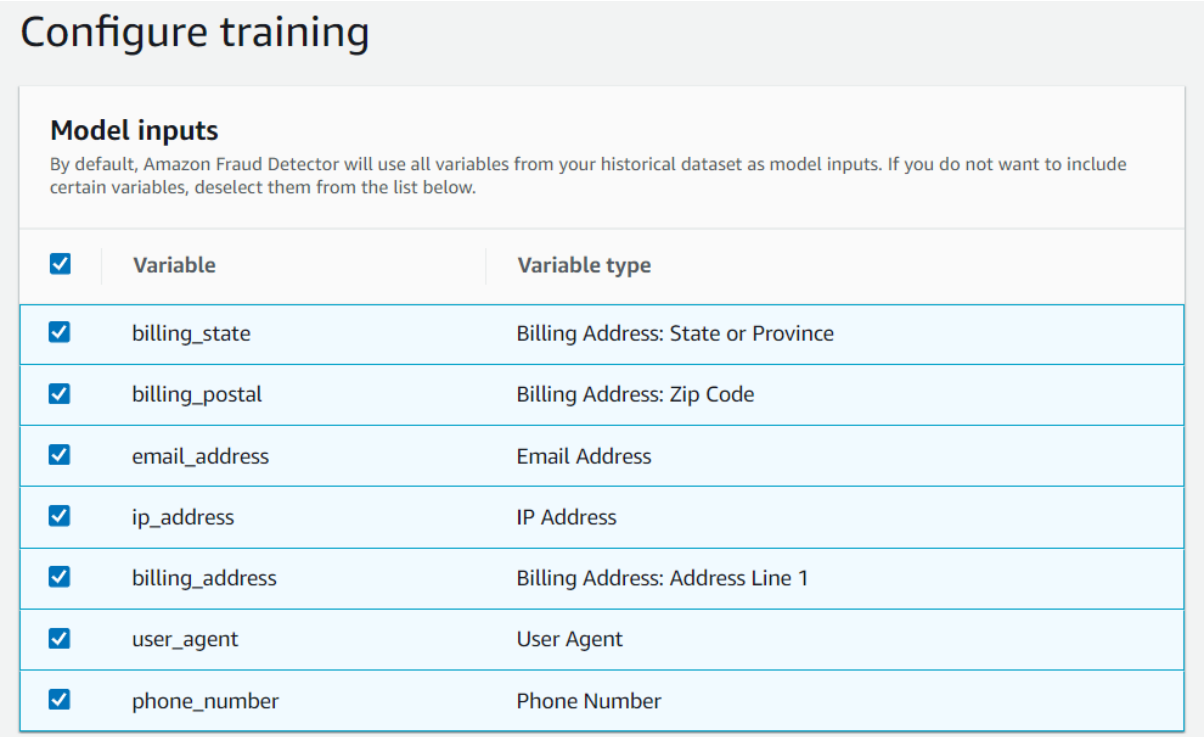
16. You need to create this role in the similar manner as you did before, give it your S3 bucket name and click on create role.



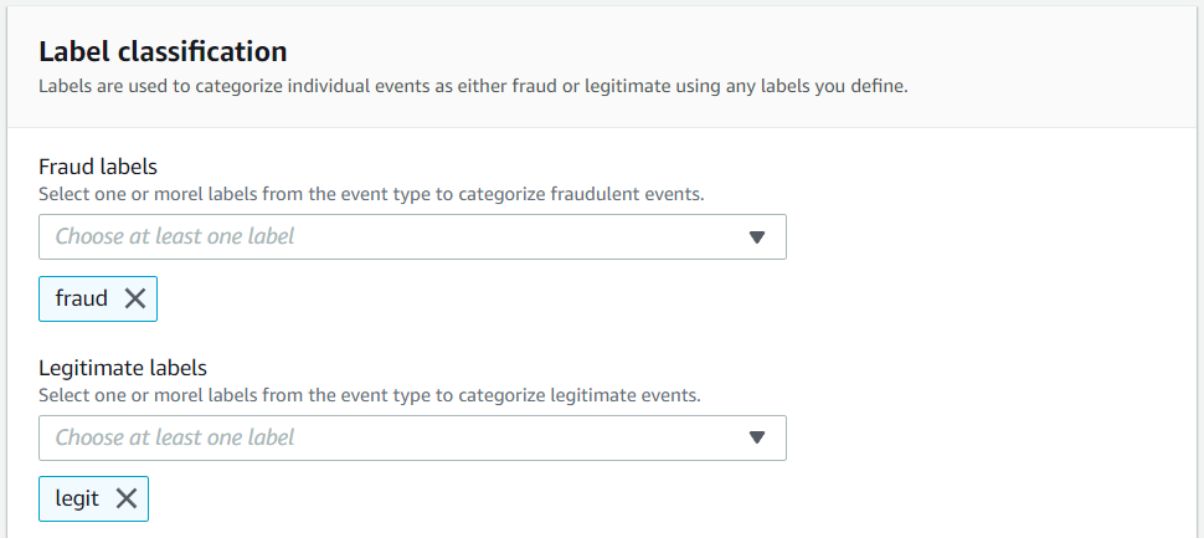
17. In the data location you need to give the S3 object URI then click on Next.



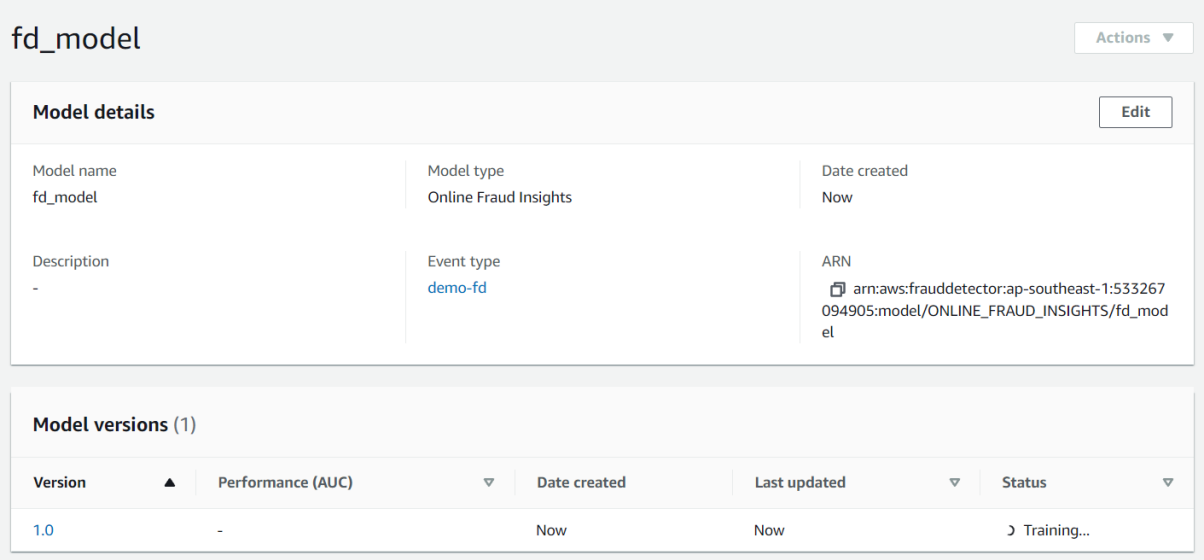
18. Here you can see that you the model inputs.



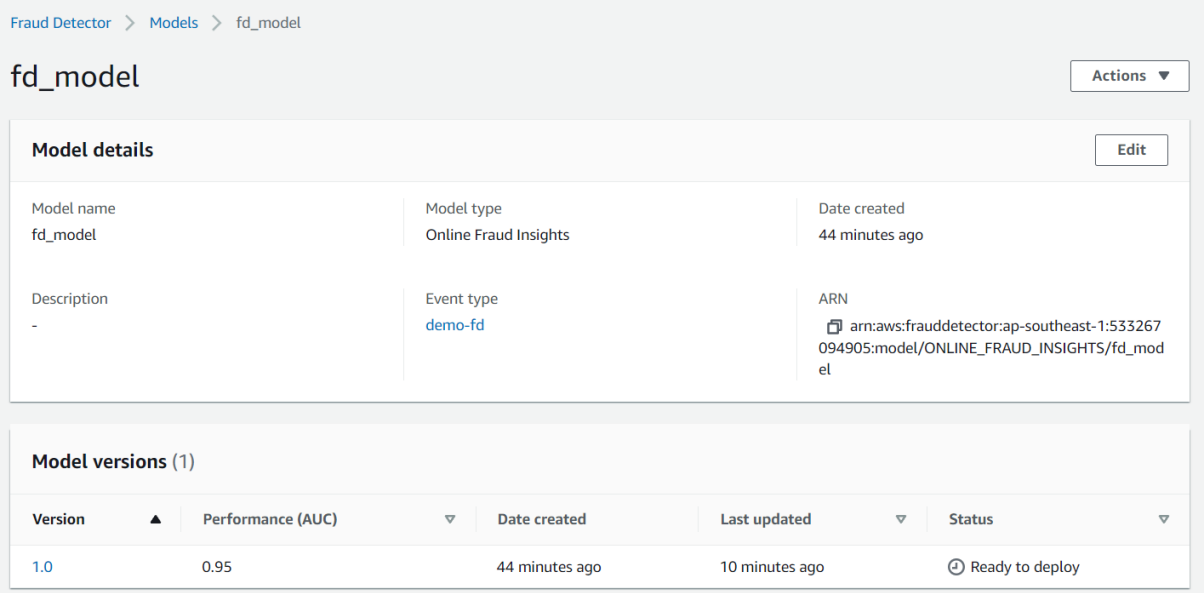
19. Then in the label classification choose your labels and move to review page. Click on Create and train model.



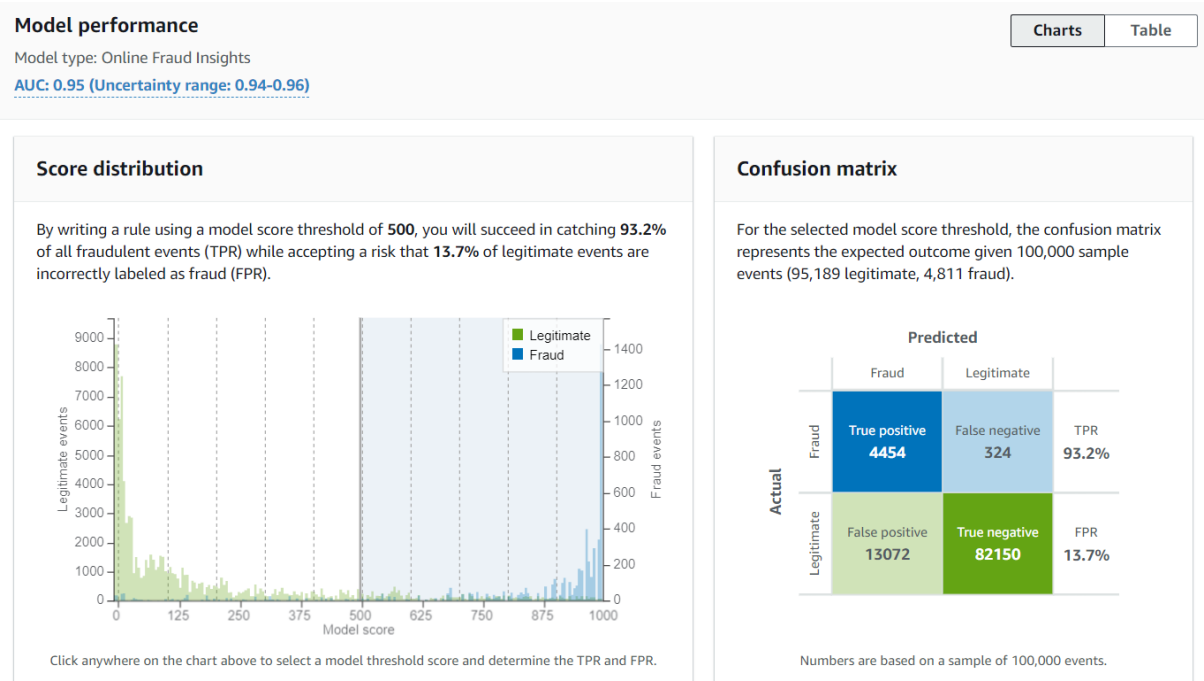
20. Below you can see that our model has been created and currently it is being trained and it will take a lot of time so you need to wait for it.



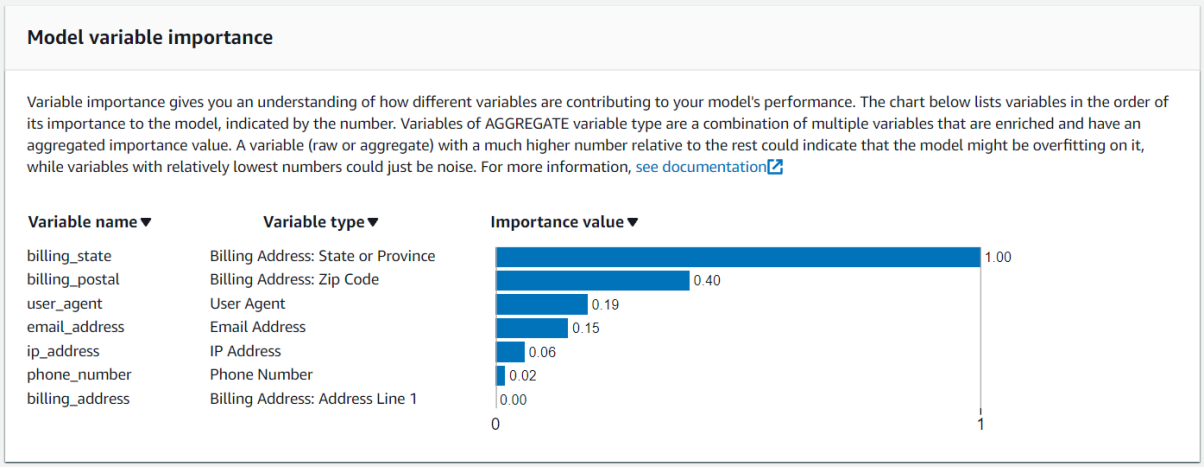
21. Below you can see that our model has been trained and it is ready to deploy. If you click on version to open it.



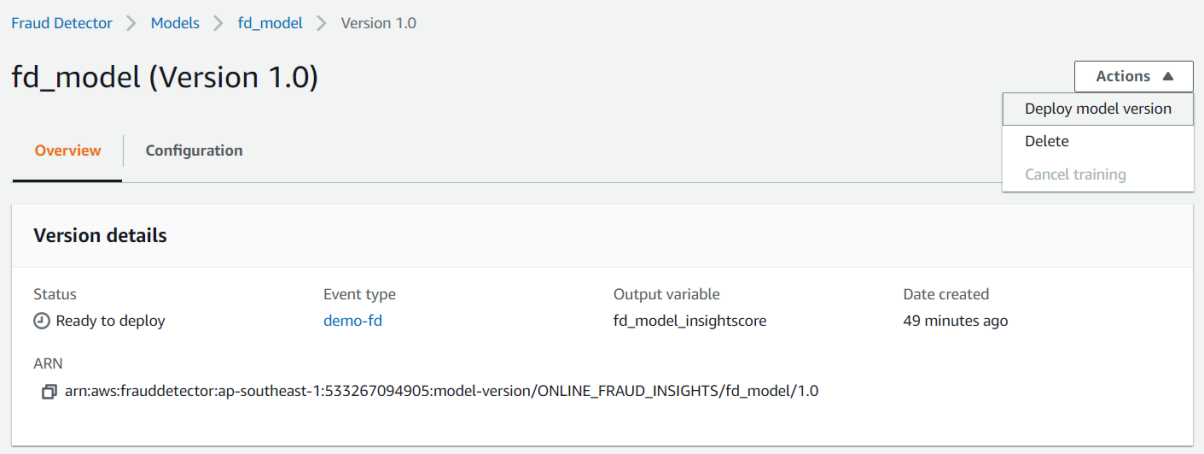
22. In this model version you can see the model performance.



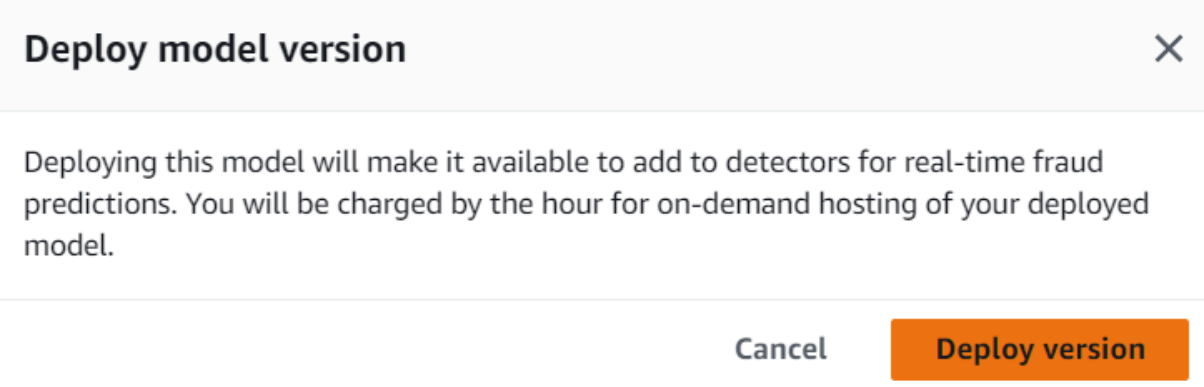
23. The model variable importance and other things.



24. Now scroll up and click on Actions and choose to deploy model version.

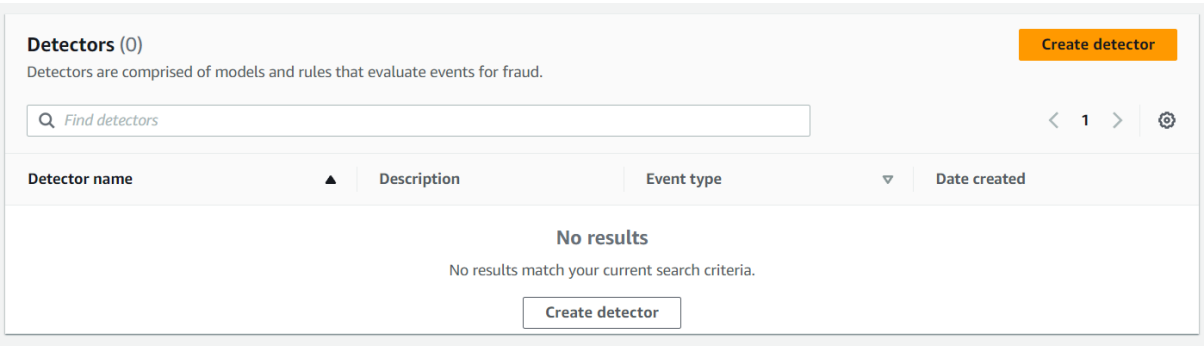


25. Click on deploy version and it will take at least 10 minutes to deploy this model.

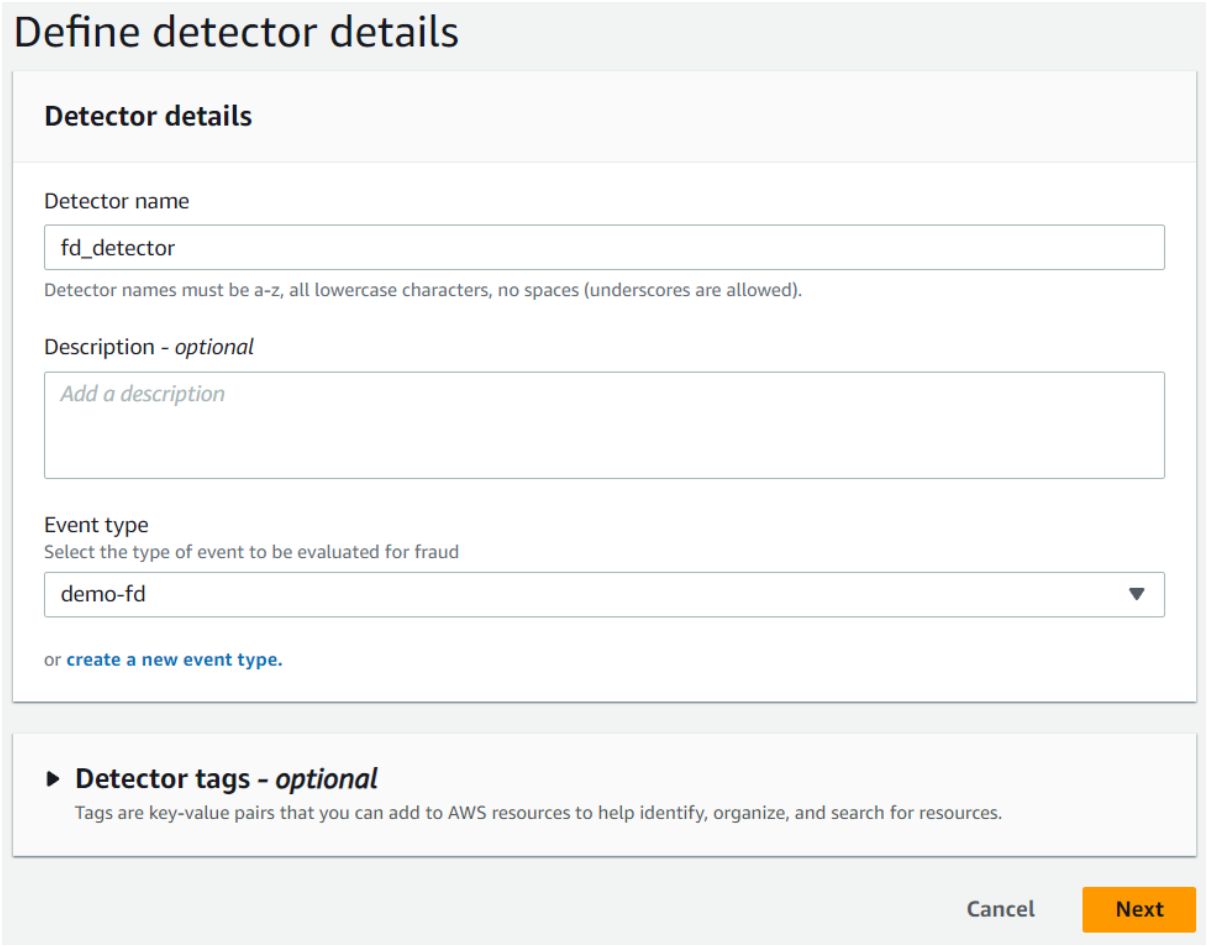


26. Below you can see our model is active now.

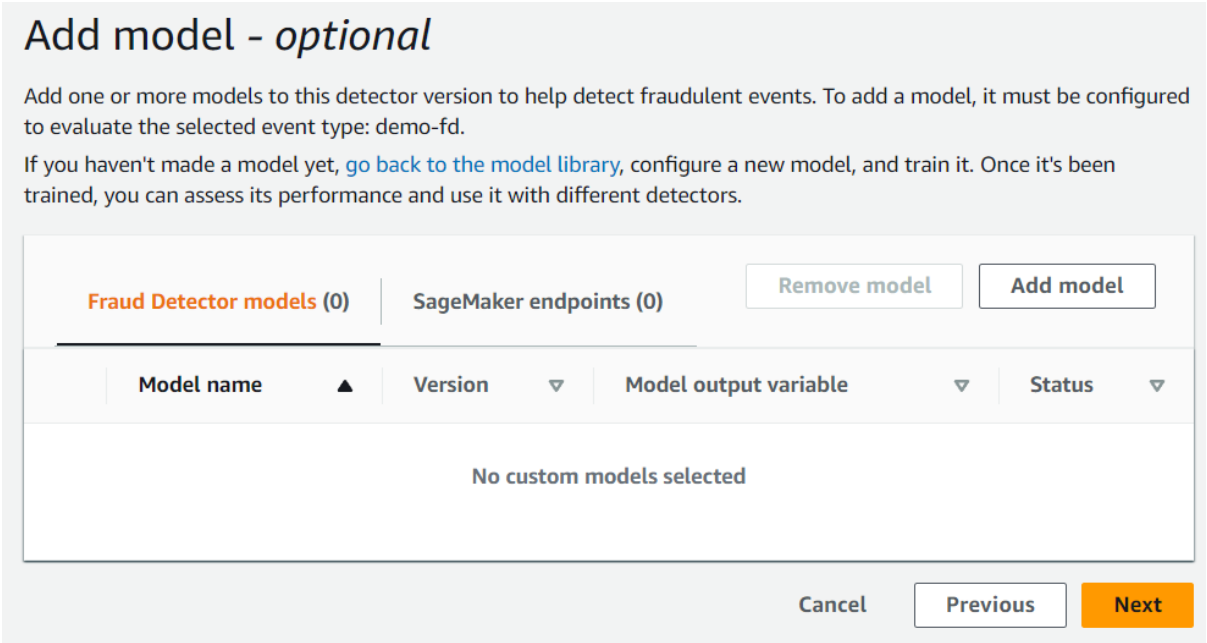
27. Now we need to create a detector. Click on Create detector.

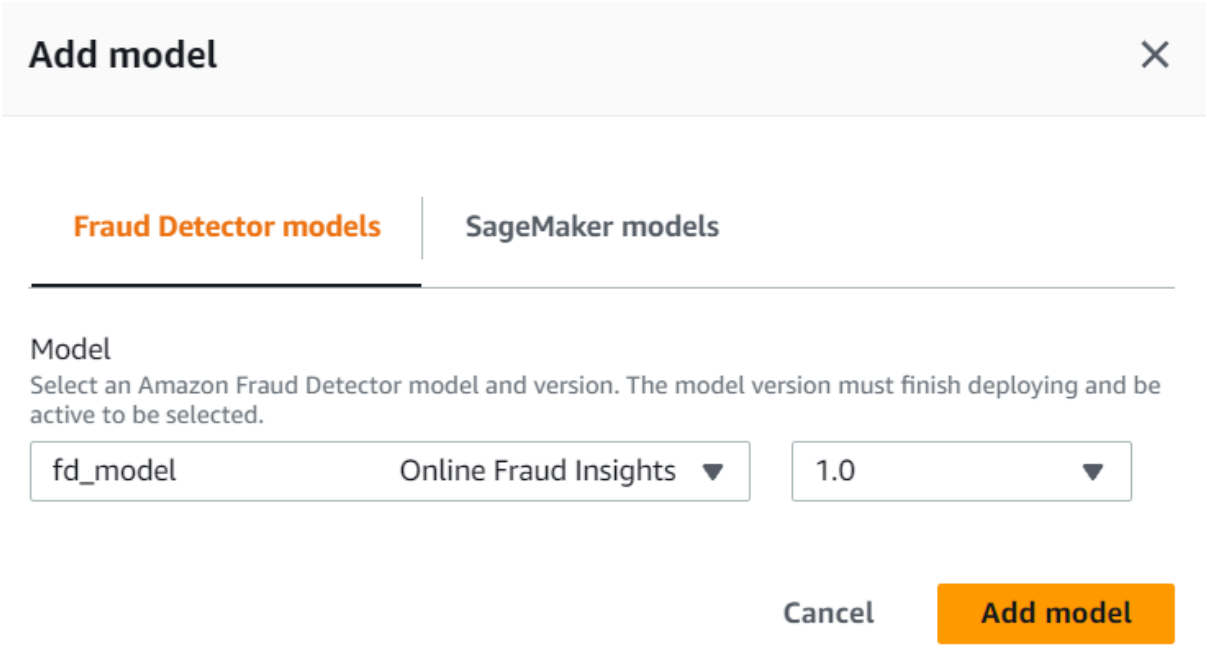


28. First you need to give it a name, then choose your event type and click on next.

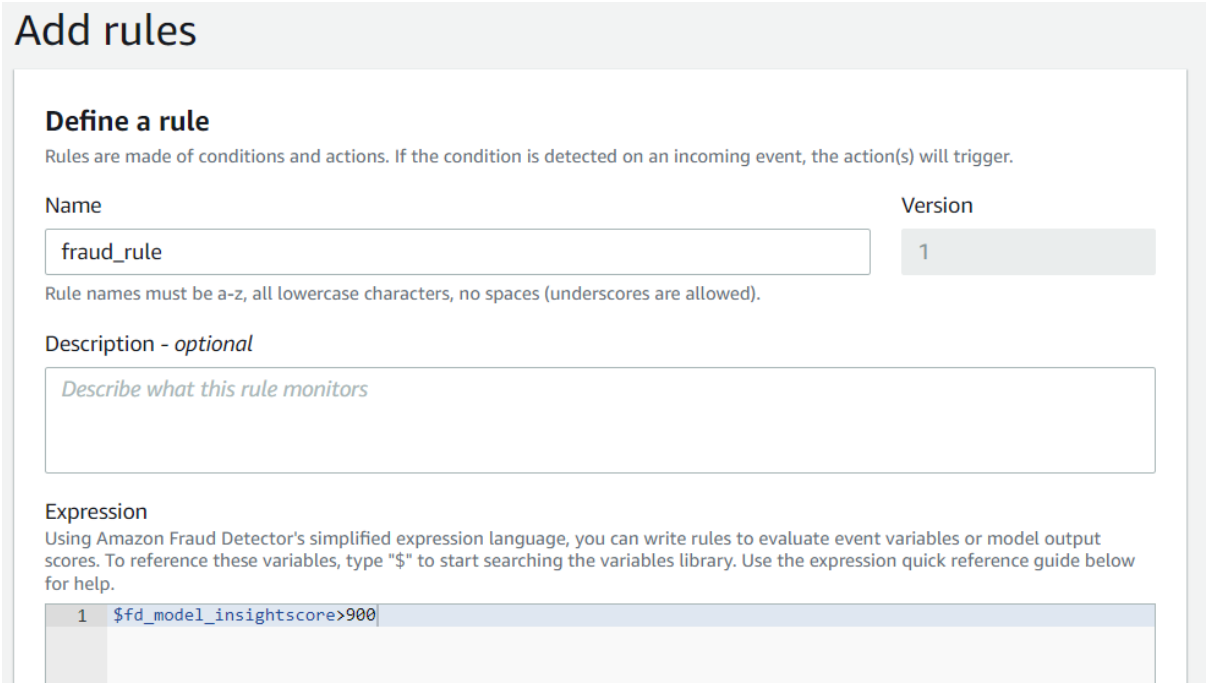


29. On the next page you need to click on add model and choose you model and version number. Then click on next.

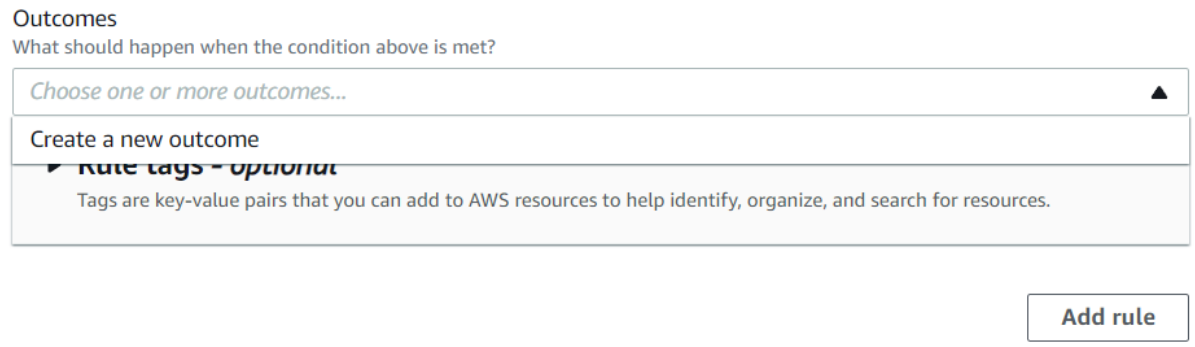


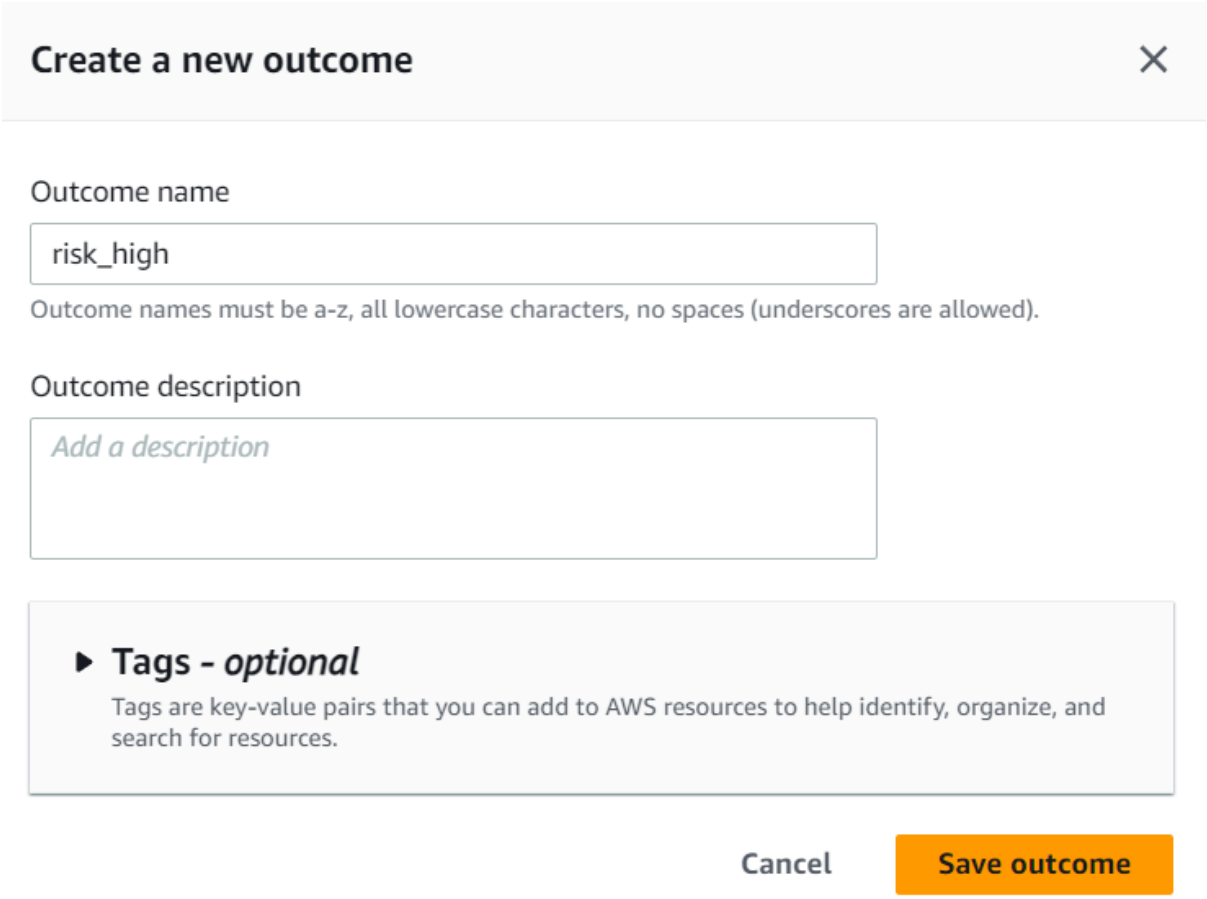


30. In the add rules section you need to write a name for your rule and in the expression, you need to give the same expression as you can see below. You just need to write $ sign and you will have the expression.

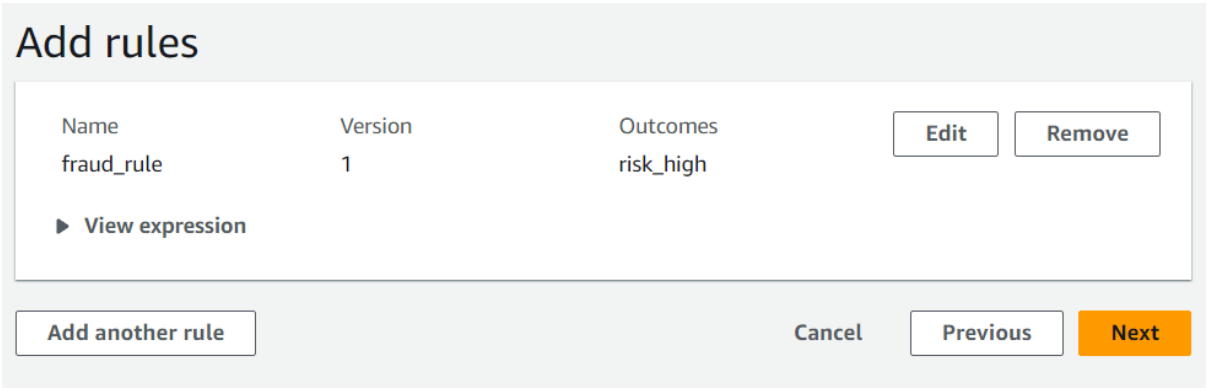


31. Then you need to create a new outcome. For the outcome name write risk_high and click on save outcome. In the end click on add rule.

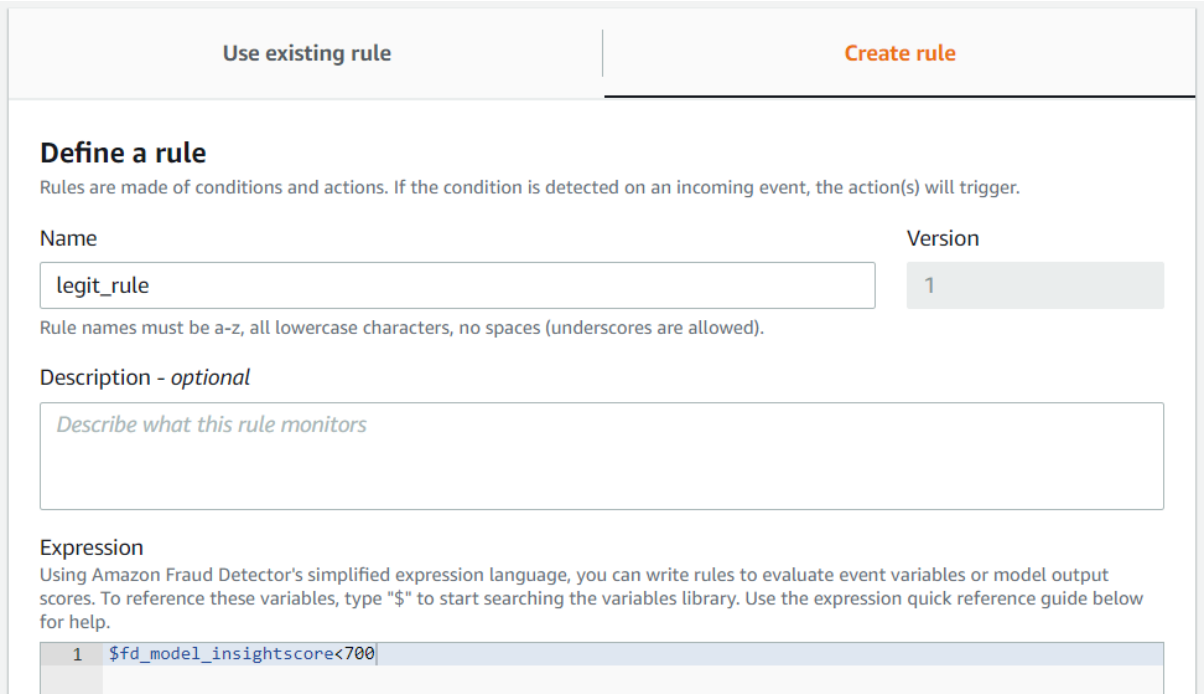




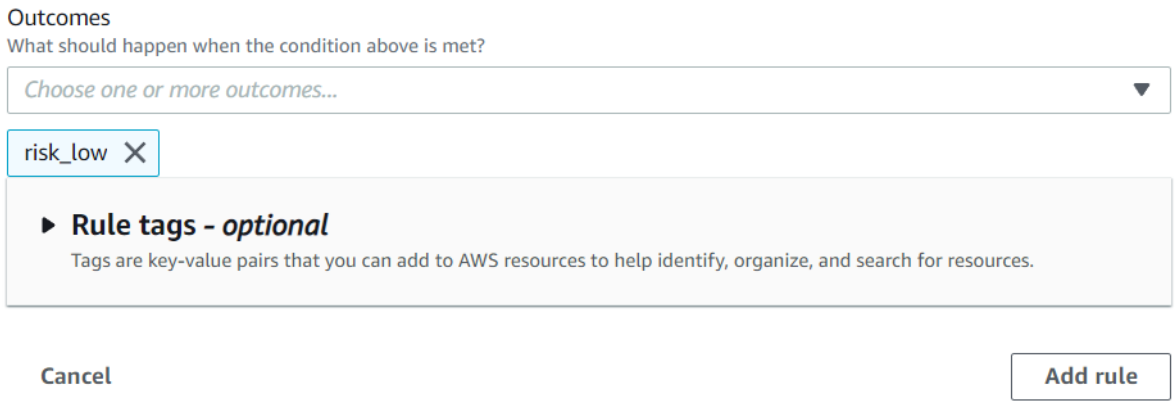
32. Then you need to click on add another rule.



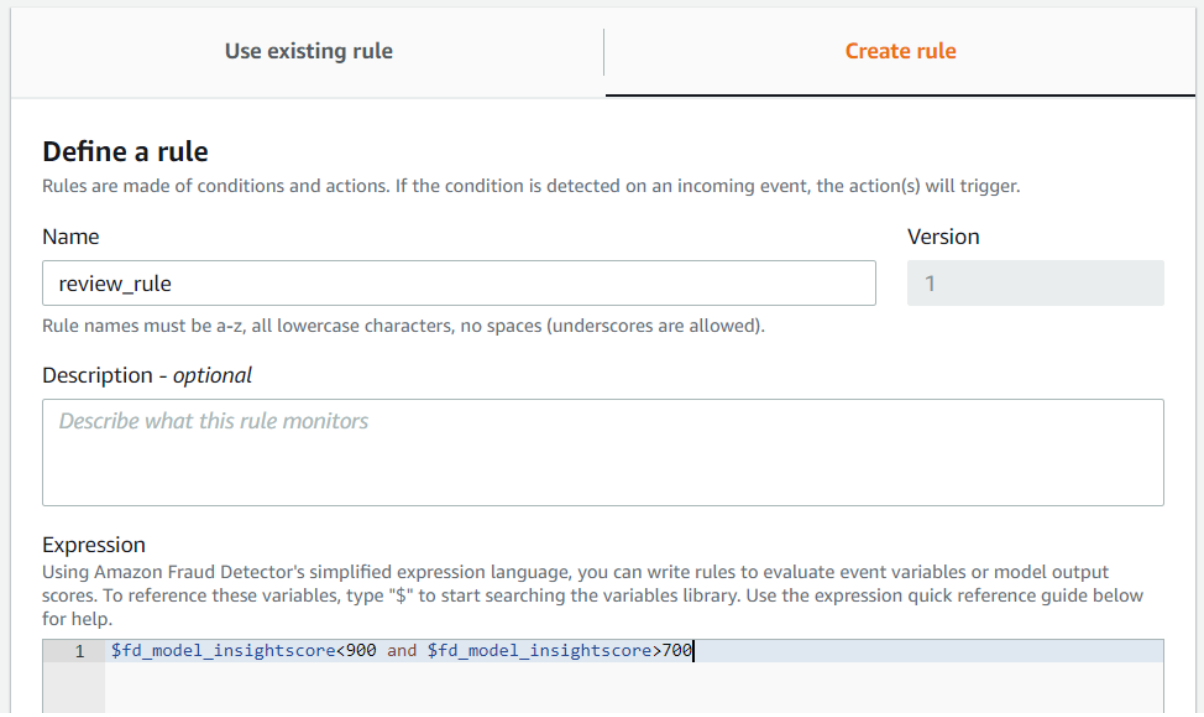
33. Choose create rule and give your rule a name and in the expression give the same as shown below.



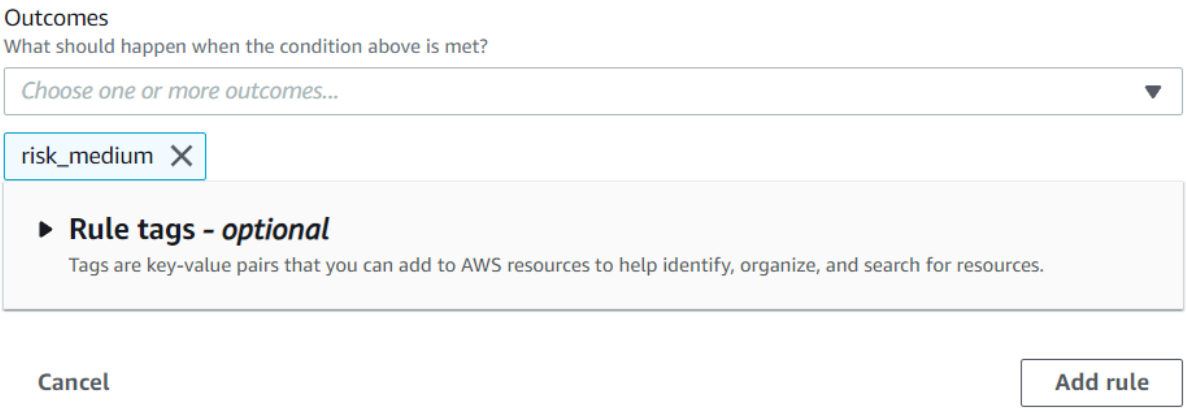
34. In the outcome again you need to create a new outcome and this time give the outcome name as risk low. Then click on Add rule.



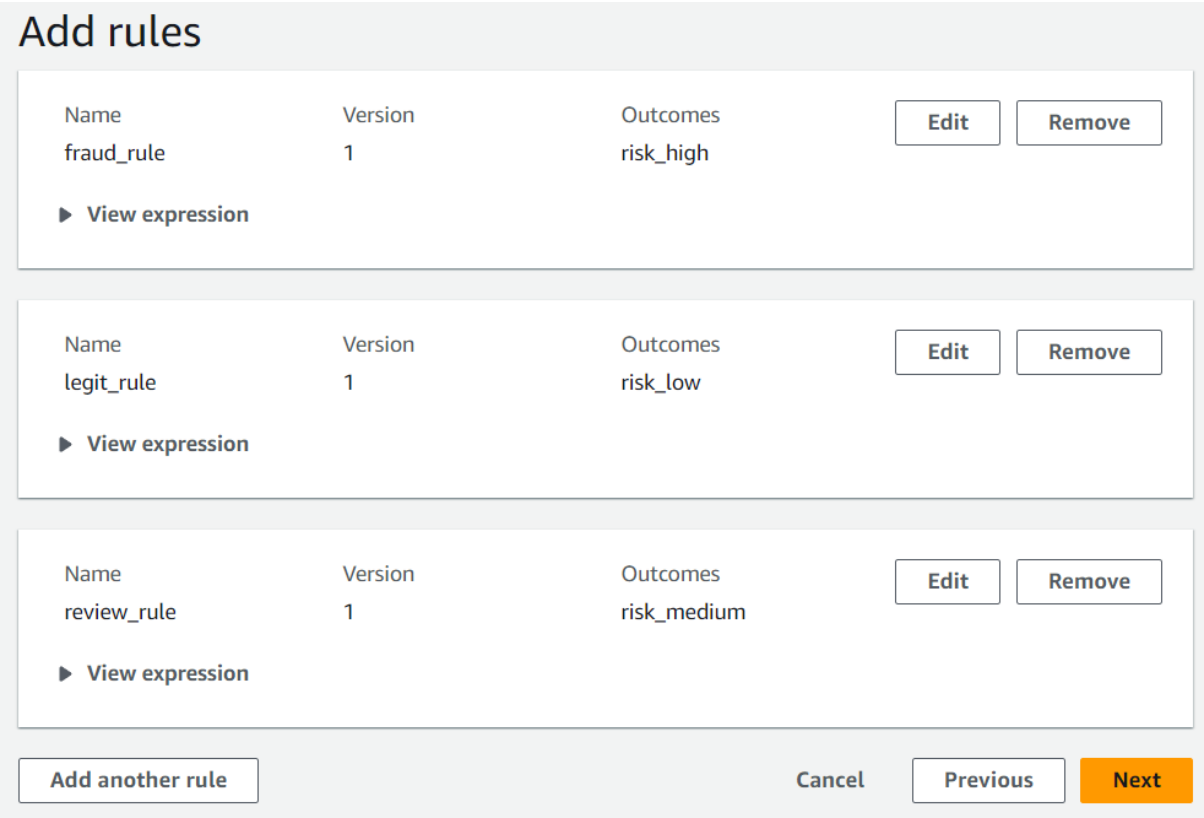
35. So, we have created two rules one is for fraud and other one is for legit. Now we are going to create another rule for review. So, click on add another rule.
36. Choose to create a new rule give it a name and then give the same expression as shown below.



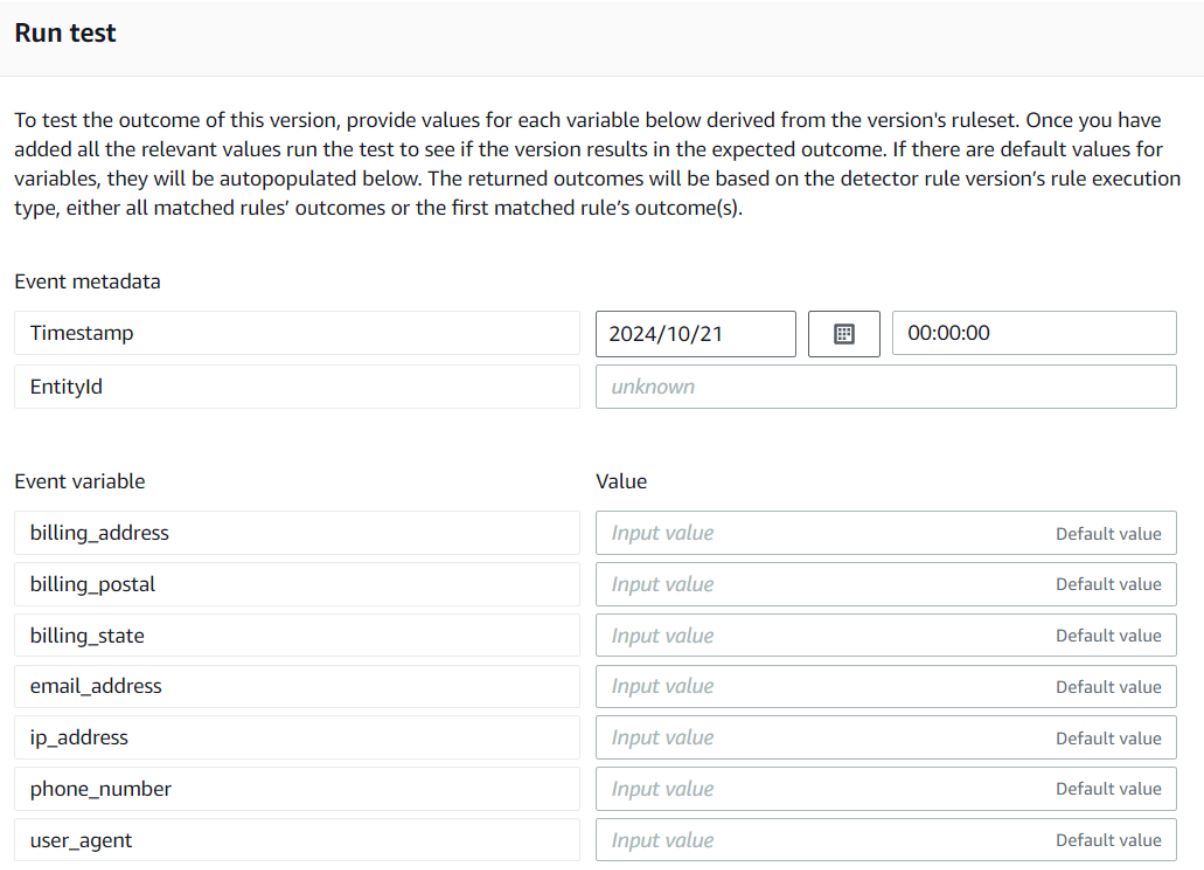
37. Then in the outcomes you need to create a new outcome as shown below and click on add rule.



38. Below you can see that we have 3 rules added. Now click on next and move to review page and create your detector.



39. We need to give the values in the run test area so that we can test our dataset.



40. Here you can see that we have filled all the values. And you can get these values from the CSV file you download and uploaded to your S3 bucket at the start of this lab.
41. Now click on run test.

Event variable	Value
billing_address	12351 Amanda Knolls Fake St.
billing_postal	34491
billing_state	NC
email_address	fake_cgonzales@example.net
ip_address	112.136.132.151
phone_number	(555)333 - 9246
user_agent	Mozilla/5.0 (iPad; CPU iPad OS 10_3_3 like Mac OS X) Apple

Run test

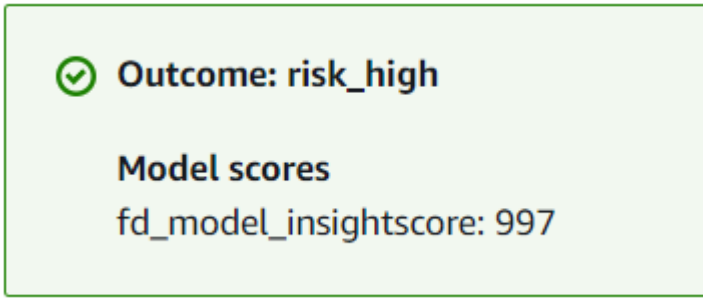
42. Below you can see that we get the outcome as risk low which means that it is legit event and its score is 153 which is less then 700.

Run test

43. Now we have taken the information for a fraud data and filled it here.

Event variable	Value
billing_address	65898 Amy Estate Fake St.
billing_postal	32931
billing_state	AZ
email_address	fake_timothysmith@example.com
ip_address	59.157.144.1
phone_number	(555)596 - 5579
user_agent	Mozilla/5.0 (iPod; U; CPU iPhone OS 3_0 like Mac OS X; sid-

44. And here you can see that we get the outcome as risk high which means that it is a fraud event whose event score is 997 which is more then 900 of value.



45. Once you are done with this lab just delete all the things.