Step by step POC process

Creating a Proof of Concept (PoC) for a loan underwriting AI agent on Google Cloud is a strategic way to validate the technology, demonstrate its value, and get stakeholder buy-in without a massive initial investment. A successful PoC is focused, has clear objectives, and proves the feasibility of the core concept. Here's a step-by-step guide on how to create one:

Step 1: Define the Scope and Success Criteria

A PoC should be narrow and focused. Trying to automate the entire loan process at once is a recipe for failure.

- **Choose a Specific Use Case**: Instead of building a full underwriting system, focus on a single, high-impact part of the process. For example:
    - **Document Processing**: Automate the extraction of key data (e.g., income, assets, debts) from a specific type of document, like a W-2 or a bank statement.
    - **Risk Assessment**: Build a model that takes a set of structured data (e.g., credit score, debt-to-income ratio) and predicts the risk of default.
    - **Fraud Detection**: Focus on a single type of fraud, such as identifying inconsistent addresses or duplicated information across a small sample of applications.
- **Define Clear, Measurable KPIs**: How will you know if the PoC is successful? Establish metrics upfront.
    - **Accuracy**: "The AI agent must extract key data from W-2 forms with at least 95% accuracy."
    - **Efficiency**: "The time to process one loan application document will be reduced from 30 minutes to under 5 minutes."
    - **Performance**: "The risk prediction model's accuracy on the test data set must be at least 80%."

Step 2: Assemble the Data

Data is the fuel for any AI project. You need a representative, clean, and a sufficiently large dataset for your chosen use case.

- **Gather a Small, Diverse Dataset**: Collect a sample of historical loan application data, including documents and corresponding final decisions. Ensure the data is de-identified to protect privacy.
- **Data Preparation**: Clean and label your data. This is often the most time-consuming part of an AI project.
    - If you're focused on document processing, you need to know what the correct extracted values are for each document (e.g., the exact income number from a pay stub).
    - If you're building a risk model, you need to clearly label the outcomes for each application (e.g., "defaulted" vs. "paid in full").

Step by step POC process

Step 3: Choose the Google Cloud Services

Select the most appropriate Google Cloud tools for your specific PoC. The goal is to use out-of-the-box services as much as possible to accelerate development.

- **For Document Processing**:
    - **Document AI for Lending**: This is the ideal tool for a PoC. It's a pre-trained model designed for financial documents, so you won't have to train a model from scratch. It can quickly demonstrate the ability to extract key fields from your sample documents.
    - **Cloud Storage**: Use this to store your raw documents and the extracted data.
- **For Risk Assessment or Fraud Detection Models**:
    - **Vertex AI AutoML**: If you have a good dataset but limited data science expertise, use AutoML. It allows you to train a custom model with a simple, no-code interface. You'll upload your structured data, define your target variable (e.g., "default"), and AutoML will build the model.
    - **BigQuery**: Store all of your structured data (applicant information, credit history, etc.) in BigQuery. It's a great platform for a PoC because it's serverless and has built-in machine learning capabilities (BigQuery ML).
- **For Orchestration and Automation**:
    - **Cloud Functions**: Use a simple Cloud Function to stitch the pieces together. For example, a function could be triggered when a new document is uploaded to Cloud Storage. This function would then call Document AI, save the results, and maybe even send a notification.

Step 4: Build and Test

- **Create the Pipeline**: Implement the workflow using the selected Google Cloud services. A simple pipeline might look like this:
    - User uploads a document to Cloud Storage.
    - Cloud Function triggers Document AI to process the document.
    - Document AI sends the extracted data back to the Cloud Function.
    - The Cloud Function stores the extracted data in BigQuery.
- **Run the PoC**: Put your test data through the pipeline. This is where you'll validate your hypotheses and test against your KPIs.
    - **Document Processing PoC**: Run your sample W-2s through Document AI and measure the accuracy of the extracted data.
    - **Risk Model PoC**: Use your trained model to make predictions on a held-out test dataset and compare the results to the actual outcomes.
- **Involve the End User**: The PoC will be more compelling if it's not just a technical exercise. Show a loan underwriter the output of the system and get their feedback. This helps identify potential issues and builds support for a larger project.

Step 5: Evaluate and Plan for the Next Steps

Step by step POC process

- **Document the Results**: Create a report that clearly outlines the objectives, the methods used, the results (measured against your KPIs), and the key findings.
- **Analyze the Outcomes**: Did the PoC meet its goals? What were the limitations? What worked and what didn't? Even if the PoC fails to meet all of its objectives, it's still a success because you've identified a problem early on.
- **Present to Stakeholders**: Share your findings with a clear recommendation:
    - **Proceed**: The PoC was a success and justifies a larger investment to build a production-ready solution.
    - **Pivot**: The PoC showed promise, but a different approach or different data is needed.
    - **Park**: The PoC failed to demonstrate value, and the project should be shelved for now.