Project Ideas

1. **MLOps as a Service:**

* **Problem:** companies today are building more machine learning models and they need a way to automate the process of converting models from research to production, monitoring the performance of models in production, continuously retraining models on new data collected, and versioning the model during the research phase. Currently, there are multiple companies creating their own suite of MLOps tools for monitoring model performance, trusted AI, etc. The problem is integrating these tools into one cohesive pipeline because these companies must:

1. Call up the vendors that are building each tool.
2. Devote engineering resources and time to evaluate each tool and decide on whether to integrate it into a pipeline.
3. If they choose to use the tool into their pipeline, now they must devote more engineering resources to integrate it into their pipeline.
4. Each vendor has their own pricing model, and the quality of their solutions are oftentimes proprietary (many of these vendors use proprietary metrics) which makes it difficult for engineering teams to truly evaluate the solution.
5. During production, the user must continue to fine-tune the parameters of these tools and update these tools to newer versions.
6. Government clients do not have access to MLOps solutions that are secure (PPML tools are almost non-existent).

* **Description:** create an MLOps as a Service platform that does the following:
  + Allow the user to choose which tool to test out and plug-and-play into their pipeline.
  + Automatically integrate the tool into the user’s MLOps pipeline upon selection.
  + Automatically set up the MLOps pipeline for the user.
  + Provides an easy-to-use interface to send data and models.
  + Standardize the pricing structure. Establish one pricing model (or a small number of pricing models) between all partnered vendors.
  + Standardize the process of evaluation between vendor and user. Have a standard set of NDAs that are signed at the beginning once user subscribes to the service so that the user can now evaluate the vendor’s proprietary algorithm/metric.
  + Manage compute resources to run these tools on behalf of the user.
  + Use the user’s MLOps pipeline and data to test our own suite of MLOps solutions.
  + Notify user of new updates, automatically runs tests to evaluate those new updates, automatically updates those tools upon user request.
* **Target markets:** 
  + Startups: low in resources and manpower and more willing to accept a solution that offloads the work of evaluation and integration.
  + Government sector: make sure tools are specifically built with security in mind. Support on-prem setup and management solutions. Create proprietary MLOps tools for PPML (e.g., model watermarking schemes).
* **Why me?**
  + **What do I currently have?** 
    - Experience: already have experience building MLOps pipelines and the tools that need to integrate into the pipeline (xRCA, FAIRS, AMA).
    - Research: conducted ML research; especially PPML research.
    - Top secret security clearance.
    - Connections: I have connections with CGI and Leidos (not sure if they want to continue working with me though).
  + **What else do I need?**
    - Connections to investors
    - Funding
    - More connections to researchers and engineers
* **Why now?**
  + **Why is two years ago too early?** MLflow and many MLOps tools are still in its infancy; much more mature ecosystem for MLOps now means more orgs are about to start using MLOps because they now can. Mojo was just released, which means a more cohesive ecosystem for machine learning development and production can now be built.
  + **Why is two years from now too late?** The US might be at war with China, making the race to productionize ML that much faster.
* **How do we get customers?** Market to startups (especially defense startups) and government defense contractors.
* **How can we sell to customers?** The sales pitch looks good thus far.
* **How can we get our product into customer’s hands?** The API must be good.
* **What sets us apart from the competition?** Databricks is building something like this last year, but I suspect they don’t have much focus on the public sector.
* **Technical skills required:** machine learning research, MLOps, cloud computing.

1. **Smart contract simulation:**

* **Problem:** smart contract developers today are building smart contracts that will later be set in stone on the blockchain on intuition and experience alone. This process needs to be much more data driven.
* **Description:** given a smart contract and information about the parties participating in the contract, develop a way to simulate how the parties will behave in a smart contract.
  + Train a machine learning model to simulate how a party behaves in the smart contract (one model for each party).
  + Use game theory to help with creating the simulation.
  + Use the smart contract to conduct reinforcement learning. Each party tries to optimize their own metric, i.e., party A wants to get more coins and party B wants to get more energy.
  + Things needed for creating a prototype: 1) need to find a situation where a simple smart contract has already been created, 2) number of parties participating in this smart contract needs to be small, 3) need to have data on what happened when the smart contract parameters changed, and 4) need to have data on how the parties respond those changes.
* **Target markets:**
* **Why me?**
  + **What do I currently have?** 
    - Experience: a decent understanding of ZKP and building machine learning models.
    - Connections: Shadman and his network of web3 people
  + **What else do I need?**
    - More experience needed with working with ZKP libraries.
    - More experience needed with multi-agent systems.
    - More experience needed with blockchain, smart contracts, and web3 development.
    - More connections with economics researchers
  + **Why now?**
  + **Why is two years ago too early?**
  + **Why is two years from now too late?**
* **How do we get customers?**
* **How can we sell to customers?**
* **How can we get our product into customer’s hands?**
* **What sets us apart from the competition?** Nobody else has this solution.
* **Technical skills required:** 
  + Blockchain, smart contracts, and web3 development
  + Game theory, economics
  + Multi-agent systems, reinforcement learning

1. AR on guns and shoot house:

* Create AR simulations for training purposes. AR simulations can include generating scenes of hostage situations involving relatives, specific targets of interests, large scale situations like people running everywhere.
* Create attachments to allow track where and when the gun is shooting.
* Have instructors develop custom tactics with drones and sell courses to teach those tactics.
* Build a miniature shoot house and CQB experience.
* Build capabilities to use Lidar sensor (available on iPhone and MacBook) to scan environment.
* Create ability to set up training environment according to the lidar information sent in by the client.
* Client trains with BB guns to conduct force-on-force training.
* Create service where we build BB guns (or simunition) according to the user’s weapon of choice.
* **How do we get customers?** Create a shoot house experience where newcomers could come in and have fun.
* **How can we sell to customers?** Record videos of newcomers getting shot and have instructors explain to them what they should have done differently and show them what other services the facility offers.

1. Compiler project:

* Implement an LLVM backend for either Donald Knuth’s MIX architecture or Cairo ZKP
* Emulate MIX on Qemu (no need to emulate Cairo)
* Implement a toy programming language that compiles down to LLVM IR