**Investigating possibilities of technology to complement Protégé Ventures’ investment approach**

1. Introduction to Technology in VC

* General statistics for why we should adopt it
* Highlight possible trends for technology in VC
* Highlight current problems and challenges that may be resolved by technology

1. Introduction to Technologies – Clarify difference between the different disciplines

* Data Analytics
* Machine Learning
* Automation Tools
* Others (etc Data Engineering, etc, TBC)

1. Description of Technologies – Specific methods within these different disciplines

* Examples in Data Analytics (i.e. Network Analysis)
* Examples in Machine Learning (i.e. Predictive Analytics)
* Examples in Automation Tools
* Examples in Data Mining, etc

1. Specific Case Studies of VCs adopting technology

* 3 Case Studies

1. Specific Case Studies of startups disrupting the VC space

* 2 Case Studies (i.e. Affinity)

1. Possible technology adoptions for VC funds in SEA
2. Possible technology adoptions for Protégé Ventures

* Steps for Data Implementation
* Determine what kind of data you want to keep track of
* Keep details of company records (Its CAC, LTV, etc)

**Current Taskings:**

1. Look to interview VCs (preferably SG) that have data science arms for more information

|  |  |  |
| --- | --- | --- |
| Vertex Ventures | Declined | <https://www.linkedin.com/in/ying-jie-tan-24486959/> |
| Arbor Ventures | - |  |
| Openspace Ventures | - | <https://www.linkedin.com/in/zhao-yiliang/> |
| Temasek Holdings |  |  |
| Hatcher+ |  | DY |
| Dymon Asia Ventures |  |  |
| Rocketship.co | USA-based |  |
| 500 Startups | USA-based |  |

1. Introduction to technology in VC – Member #2
2. Research on different forms of technologies – Member #3
3. Research on specific VC Case Studies online and/or reach out for more information

* CASE STUDY 1 – Member #4
* CASE STUDY 1 – Member #5
* CASE STUDY 1 – Member #6

<https://medium.com/hackernoon/winning-by-eating-their-own-dogs-food-83-venture-capital-firms-using-data-ai-proprietary-da92b81b85ef>

1. Research on specific startups in the dealflow space and/or reach out for more information

* Affinity > Relationship Intelligence – Member #7
* (STARTUP #2) – Member #8

|  |  |
| --- | --- |
|  | Taskings (if any): |
| Nicholas Koh |  |
| Augustine Thia |  |
| Gordon Ng |  |
| Tee Zhi Zhang |  |
| Marcus Wan |  |
| Benedict Chong |  |
| Kelvin Sutedja |  |
| DY | Reaching out to people NOW |

**GUIDING QUESTIONS (INSERT MORE QUESTIONS OR ANSWERS / RESEARCH HERE)**

**INTRODUCTION TO TECHNOLOGY IN VC**

* How does technology help VC?
* What are some statistics that can show the impact of Data Science?

**INTRODUCTION TO TECHNOLOGIES**

* What are the different forms of technologies?
* Explain difference between data analytics and machine learning, data engineering, etc
* What are the different forms of machine learning? (Neural Networks, different types of NNs, text mining, regression for probability of success, etc)

**RESEARCH ON SPECIFIC VC FUNDS AND DATA SCIENCE**

* Some online case study research
* Conduct possible interviews with VC funds locally that have data science arms

**RESEARCH ON SPECIFIC CASE STUDIES OF STARTUPS**

* Some online case study research on startups
* Conduct possible interviews with startups that are working in this space

**OUR THESIS FOR HOW VC FUNDS CAN ADOPT DATA SCIENCE**

* What are some technologies that every fund can adopt?

**OUR THESIS FOR PROTÉGÉ VENTURES**

* **­**Do we consider the option?
* What are some immediate wins we can tackle
* What are some long-term wins we can plan for
* What are some options we cannot possibly do understanding the context of PV

**INTERVIEW QUESTIONS WITH DR ZHAO YI LIANG:**

Q: What are some applications of data science and analytics in the VC space?

Q: Is data science applicable for every single VC?

* Not relevant for very early stage VC funds
* Cast net very wide and hope for the best

Q: How does data science help in fundraising, investments or operations?

Q: Would you know any other data scientists in the VC space that we could approach?

Q: In what ways has data science helped in dealflow and sourcing of deals?

* Determine which startups the VC should identify first

Q: Given that VC is largely a people-oriented business, and there is a lack of hard concrete data about startups, how does the role of a data scientist differ from that of the regular venture analyst?

* What kind of information that you want to keep track
* Details of company performance (Revenue, Number of Customers…)

Features

Output:

* Things to predict

Build a binary classification model:

* Yes or No results
* You need training data
* What companies you have spoken with

What is successful?

* Whether or not they valuation above certain number
* Upround
* IPO, M&A
* What other factors they succeed

With definition of success, we can predict startups

* Binary classification can give probability of startup, whether or not it will succeed.

Challenges:

* Availability of data, and variety of data
* Different industries have different data points and different focus
* SaaS, B2B, etc have different forms of data points
* You need to build multiple classification models
* You can use webtraffic to obtain data points
* Model tuning and feature engineering also takes time
* Subsequent evaluation takes a lot of time, and model development takes a lot of time

Companies that are seeking investment we can get data points from them as well.

Other than that, we also look for data providers.

Graph based models. Some VCs follow founders instead of opportunities. Network analysis, network graph embedding, develop feature generation from graph structure.

Data Science (Portfolio Support):

* Make sure they can proceed with subsequent rounds and exits
* Help portfolio companies come up with data science roadmaps
* How big data tunes should be, and of hiring data scientists
* Work together with the tech team of these startups on a technical level
* How the downstream applications should use and determine success levels
* Differentiating factor in helping portfolio companies
* Most important task (70 – 80%) portfolio support, working with 6 – 8 companies
* 15% do due diligence and aggressively grill the startups on their tech
* Other 15% to work on other projects

**Data Science at Vertex:**

Removing chunks of text:

* Unsupervised learning (NMF), via clustering
* Based on text input, come up with classification
* Using NMF, and neural network based off LSTM

Ensemble Boost:

* XGBoost to test the accuracy of the model
* LSTM NN best performing model apparently, not surprising though

NMF:

* Pick up words from texts and slide deck, Classifier is One vs Rest, naïve bayes and log regression classifiers are not so good
* Goal: Dimensionality Reduction (reduce attributes of the dataset)