

Task	Sub Task	Dependency	Owner	Status	Due date	Notes/Issues	
Capstone Schedule - Task List			Thong	Done	10/4	Ongoing updates	
Weekly Meeting Not	https://docs.google.com/document/d/1IUTJGo2NP8Zyq8PR-mjt4UX-_vdo74YvDBApfbOP9oc/edit#		Thong	Ongoing		Ongoing updates	
Workshop 1							
	Scope		Thong	Done	10/3		
	Schedule		Sarah	Done	10/3		
Presentation 1					10/10		
	Draft Outline - slide 1-3		Sarah	Done		Suggested Division of slides: Sarah (1-3);	
	Slides 4-6		Zhongqiao	Done			
	Slides 7-8		Shankar	Done			
	Slides 9-10		Thong	Done			
	Get images and beautify		Thong	Done			
	Post it		Thong	Done			
Project Description to post on wall					10/10		
	First Draft		Sarah	Done			
	Edits		Thong, Shankar, Zhongqiao	Done			
	Post		Thong	Done			
Presentation 2					11/9		
	Presentation Outline		Sarah	Done			
	First Draft		All	Done			
	Get images and beautify		Thong	Done			
	Post		Thong	Done			
1) Read ticker data for S&P stocks from Yahoo Finance			Thong	Done on 10/6	10/6		
	Full data dump (start date = 2000-01-01)			100%			
	Incremental data dump			100%			
	Unit test			100%			
	Check-in code, data			100%			
2) Read data from Twitter			Sarah	Done	10/6	Not very useful because not histoical; can work on building in for real time updates	
	Current and historical data?			100%			
	Unit test			100%			
	Check-in code, data			100%			
3) Read historical news			Zhongqiao	Done	10/8	Only got data since 2010. Will try google API	
	Full data dump (start date = 2015?)			90%		Google api can get back as far as 2015	
	Incremental data dump			90%			
	Flag: Q1, ... Q4, FY		Zhongqiao, Sarah		11/12	Not feasible because news data don't always contain quarterly info	
4) Read revenue data about a company			Shankar	Done	10/8		
	static data dump: company size, employees				10/12	Comprehensive Data available via API from Datanyze or FullContact. But only 100 API calls per month are free, basic subscription fee of \$99 per month allows up to 2500 calls per month	
	300 stocks		Shankar	Done	11/17		
	Dividend, PE ratio (not historical), revenues (historical?)		Shankar, Thong	Done	10/8	Revenue and Dividends done, searching for FREE historical PE ratio.	
				complete for b'Date, Open,High,Low,Close, Volume,Dividend,Split, Adj_Open,Adj_High, Adj_Low,Adj_Close, Adj_Volume'n' based upon stock symbol			
	Historical data - Full (start date = 2008)				10/8	Thong: where is the sp500 data? : No free data available. We can try to get data for \$40 - \$99 per month.	
						Have not found any free sites that offer Historical data	
					10/8	Basic client committed, works on specific ticker symbol, more development will happen in coming weeks	

Task	Sub Task	Dependency	Owner	Status	Due date	Notes/Issues	
5) NLP on news				Done	11/6	Must be done sooner than 11/10 to unblock baseline, final model	
	sample output data		Zhongqiao	Done			
	Remaining data (340)		Zhongqiao	Done	11/18		
	Generate sentiment analysis scores		Zhongqiao	Done			
	Finalize feature extraction		Sarah		11/24	Basic features for MVP, better features to come for next round	
6) UI (MVP)					11/10		
	Basic Wireframes		Sarah	Done	10/23		
	Review Wireframes		team	Done			
	Basic Flask Apps		Sarah	Done	11/6	MVP	
	Code for basic flask app up on github (with local-ly runnable html files)		Sarah	Done	11/12		
	Static HTML pages for main page explaining what and risks	None	Zhongqiao		TDB		
	Basic formatting on interactive page	Code on github	Sarah	Done	TBD		
	Connect the flask to back end	8,9	Thong	Done	11/20		
	Finished Flask Apps		Sarah/Thong	Done	12/01		
	Improving UI		Thong	50%			
	Deploy App to AWS		Sarah	Done			
	Automate git pull from AWS		Sarah	50%			
7) Baseline Model							
	Build training and test data (just stock prices)		Thong	Done	10/15		
	basic model with just price data	1	Thong	Done	10/20	The current model is not good at prediction stock price. Will have to try others	
	normalized data	1	Thong	Done	10/30		
	Unit test		Thong	Done	10/30		
	Check-in code		Thong	Done	10/30		
8) Final Model					12/6		
	LSTM model with just price data		Thong	Done			
	price + other financial + NLP news data	1,2,3, 5(?)	Thong	Done		This is the most laborious step to integrate and rerun everything	
	identify stocks with highest returns		Thong	Done			
	train & predict with 500 stocks (30, 45, 60 days)		Thong	Done			
	Data Analysis to narrow down features to use		Thong	Done			
	Tuning for overfitting		Thong	Done			
	Unit test		Thong	Done			
	Check-in code		Thong	Done			
9) Use ML output to recommend portfolio given investor goals (given risk choice,		7,8			11/30		
	Output: top 10 list of stocks		Thong	Done			
11) User evaluation	Conduct survey and create MOSCOW list		Sarah	Done	12/3		
	Remove input amount		Thong	Done			
	Display top 10 results		Thong	Done			
12) Finalized UI							
	Revamp UI for better usability		Thong	Done			
	Redo UI to serve table instead of graph		Thong	Done	4/12		
	Images of potential graphs on slack		Sarah/Shankar	Done	2/12		
	Develop background page that communicates value proposition		Sarah	40%	3/12		
	Deploy latest code to AWS		Sarah	Done	4/12		
	Get S&P500 Index fund data for 30, 45, 60 day gains		Zhongqiao	Done	3/12		
	Display the fund's avg gains on the UI	Data from index fund	Thong	Done			
	Display the fund as baseline in UI	Data from index fund	Thong	Done			
	Implement improved graph		Sarah	Done	3/12		

Task	Sub Task	Dependency	Owner	Status	Due date	Notes/Issues	
	Second Round of UI testing	Updated UI	Sarah	first tests of background text done; full user testing appointment scheduled weekend of 9 and 10	10/12		
Nice to haves:							
	Improve graphic design of main page	Other main UI changes	Shankar?/Sarah				
	Show more information about companies						