## Kick-off Notes

The Trade Perk
Tiefei Ma & Chris Hanker.
Tracking Brand Sentiment via Public News.
Ads allow uses to use the internet for free
1000 and was 10 we me menter you gree
Opposite is a paywall.
opposite le la gargiani.
h
Ado work woning Open Real-time Bidding.
His a knid of anction for ads.
SIT -> Jener hosting the ad. Serow note. Matform.
SSP -> Sener hosting the ad. "Demand-orde Platform".  DSP -> Componers that bord. "Demand-orde Platform".
Timerant after 100 ms.
<b>,</b>
~ 30 millioni audiani per second.
·
DMP - to "Demain Management Platform"  Kaups track of the users demographen  which is given to the SSP/DSPs.
Keeps track at the users demographen
which is onen to the SEP/DEPC.
Can't violate user privacy information. UPI PII - Personally Identifiable Information.
PTT - Personally Idea to have how how
1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -

Often componers herd to advertise to change public sentiment
Often componus herd to advertisé to change public sentiment after a blunder.
Can use NLP, etc. to develop an algorithm for tracking brand sentiment
brand fentiment
Prpelme
· Figure
O AT Applie
 AI Analytica
Can be as simple as a scale, could be multi-babel?
Challenger.
- How to extract new article: content from web page?  - How to identify acticles mentioning a certain brand?  - How to detect centiment in a broad?  - How to do all of them at kale? - Mani Mi challenge.
- How to identify actives mentioning a contain bround?
- How to defect sentiment in a broad?
- How to do all of them at kale? - Mani Mc challenge.
Available dataset -> could use an API or Scraper
Available dataset -> could use an API or scaper  to use common crawl => commancrawl.org,  monthly web scaping since 2014.
monthly web scaping anie 2014.
monthly releases 360TB of data.
Pandes an HTML and text processed formats - Clatter isn't
actent)
What is Spark?
· · · · · · · · · · · · · · · · · · ·

NLP libraries -> John Snow Labor & Hugging Face
Contangration - Mike law and Dales
MI flow -> Platform for model deve bonnet / registry
NLP libraries -> John Snow Labo 3 Hugging Face. Contamoreation -> kubeflow and Docker  ML flow -> Platform for model development / registry.  ML Horanci -> PyTorch 3 Tensorflow.
, , , , , , , , , , , , , , , , , , ,
Use Gittlub.
MVP:
- Cloud-based data possessing and ML pipelines
- Model Evaluation woults
- Database for times series -> \$3? Mango? Or Kelational.
- Cloud-based data paressing and ML pipelines - Model Evaluation. Moulti Database for times series> 53? Mango? Or Relational Interactive Darkboard for tracking.
Pl La
 Stretch goals:
- Build our app that wer this data (can be a prototype).
Can start with top 500 companies (but should be general).
shake and : To and some some the selection
Stretch good: To give reasoning why the sentiment.  Also com measure the sentiment within sentences and rank  Them
Then -
Bace minimum for dashboard is selected company
Bare minimum for dashboard is selected company and time range
7
Can add more though
Publication is possible.
N POSTINE.
Start brain storming during holidays. Meetings nichally ad hoc
Start boarn storming during holdays. Meetings inihally ad hoc and then regular. Definitely meet before week 2 of 50 mg.

Deadlines.  4th Feb: Report 1
1.th
4 Feb : Report !
? : Demonstration.
2nd May: Final Report & Resentation.
a ray That Report & mosacionide.