***Foundations of Data Science Capstone Project Ideas***

**Idea #1- Content Recommendation Engine**

Business Problem:

As a non-profit ministry, Educational Media Foundation not only receives hundreds of thousands of prayer requests yearly, but each day their staff gathers 3 times a day to pray for these requests. Additionally, our full time pastoral staff is available for prayer and outreach 24/7, 365 days a year. EMF would like to provide some additional content recommendations when the prayer request has been submitted, which could include podcasts, articles, music, etc. This content could be either internally developed content or external content.

Associated Papers or Topics:

There are several organizations that receive prayer requests from their listeners. However, in the ministry/non-profit industries, most of these organizations just accept the prayer request and don’t offer any additional information for the submitter between the time it is submitted and followed up on. This would be a great value add for our organization. While it is nice to speak to a pastor or know that your prayer request is being prayed for, one’s spirituality and prayer life is very personal. God can use people and content in combination to provide wisdom, peace or knowledge around a specific situation.

* <http://www1.cbn.com/biblestudy/three-keys-to-answered-prayer>
* <https://www.kernersvillenc.com/wp-content/uploads/2016/05/How-does-God-speak-to-us.pdf>
* <http://www.crosswalk.com/faith/spiritual-life/8-ways-god-speaks-to-us-today.html>
* <https://healthnetwork.com/blog/how-music-heals-the-mind-body-and-soul/>
* <http://www.npr.org/2011/06/01/136859090/the-power-of-music-to-affect-the-brain>

Potential Datasets:

* EMF Internal Prayers Database: This database houses each prayer received, if it was prayed for and by whom, the category of the prayer, dates, EMF network, the current status the prayer is in (received, read, reviewed, etc.), and the source of how it was received.
* EMF Internal Content Management System
* EMF Bible Verse System
* TED Talks Database: <https://www.datazar.com/file/full/f056e5bd0-c097-4177-a77d-6e71ef4dc66d>
* Million Songs Database: <https://labrosa.ee.columbia.edu/millionsong/>

Potential Solution:

Analysis of the prayer database will provide EMF additional knowledge around a specific segment of customers and help the organization gain awareness of some of the biggest issues facing our audience. By creating a recommendation engine, we will be able to provide resources during the immediate time of need for the prayer request submitter. In the long term, this information could be analyzed to identify gaps in needs vs. current content available.

**Idea #2- Email Text Analysis**

Business Problem:

Organizations send thousands of emails every hour. Due to lack of sophisticated monitoring of large amounts of email, most organizations may have a hard time spotting potential risks such as fraud, pre-mature product leaks, etc. Additionally, customer service management may not have a way to effectively gauge satisfaction levels among customers who interact with their brands via email. Organizations need a way to understand the context of communications in large volumes.

Associated Papers or Topics:

* Enron Scandal: <https://en.wikipedia.org/wiki/Enron_scandal>
* Hilary Clinton Personal Email: <https://en.wikipedia.org/wiki/Hillary_Clinton_email_controversy>
* Text Mining: <https://forteconsultancy.wordpress.com/2010/05/17/text-mining-going-way-beyond-just-listening-to-the-voice-of-the-customer/>

Potential Dataset:

* <http://www.cs.cmu.edu/~./enron/>

Potential Solution:

Perform text analytics on a mass set of emails to determine the context of the conversations and identify potential risks in conversations.

**Idea #3-Non-Profit Segmentation and Ranking Analysis**

Business Problem:

Donors have a hundreds of choices when it comes to supporting non-profits. How do we really know if the charity we are supporting is the best one and if it does what it claims to? Through our research internally at EMF, we know many donors spend hours researching nonprofits by doing google searches, reading reviews and talking to friends before they will make a donation.

Charity Navigator is a trusted organization for evaluating specific charities and Charity Watch provides a ranking of charities. However, neither of these organizations provide a way to compare and contrast charities on a variety of variables. Just because a non-profit has a score of 90-100 or an A/A+, does that necessarily mean the majority of your donation is going to programs vs. administrative costs? How do you compare and contrast the impact of an organization. Do they serve 20k in a very specific community or thousands world-wide. Does a better rating by trusted associations equate to more donations for that non-profit?

Associated Papers or Topics:

* <https://www.charitywatch.org/top-rated-charities>
* <https://www.charitynavigator.org/>
* <http://www.consumerreports.org/charities/best-charities-for-your-donations/>
* <https://www.charitynavigator.org/index.cfm?bay=content.view&cpid=4756>

Potential Datasets:

* [**https://www.irs.gov/uac/newsroom/irs-makes-electronically-filed-form-990-data-available-in-new-format**](https://www.irs.gov/uac/newsroom/irs-makes-electronically-filed-form-990-data-available-in-new-format)
* [**https://aws.amazon.com/public-data-sets/irs-990/**](https://aws.amazon.com/public-data-sets/irs-990/)
* [**https://archive.org/details/IRS990-efile**](https://archive.org/details/IRS990-efile)
* [**https://twitter.com/licyeus/status/743308612672466944**](https://twitter.com/licyeus/status/743308612672466944)
* [**https://gist.github.com/licyeus/95b99d6feb423ebea604b5f3e2cdf590**](https://gist.github.com/licyeus/95b99d6feb423ebea604b5f3e2cdf590)

Potential Solution:

Create a ranking system that evaluates non-profits against each other based on specific variables chosen by the user (i.e. Best use of $, Most liked by Peers, Closest to me, etc.)