Meeting Minutes

Location: Conference Call

Time: 9 am – 9:30am

In attendance: Varada, Aaron, Fan, Ayla, Martin, Nasim

# New Agenda Items – 5 mins

* Try to send the agenda a day or two ahead

# Follow Up Items – 5 mins

* Timeline on full data set
  + Full speed ahead on solution
  + Put in request for IDIR accounts
    - Re-activation of Ayla’s account
    - New accounts for Aaron & Fan
    - Requester on leave today but Martin will check in shortly
    - More details coming today
    - BC Stats has set up a shared computer
      * Be able to remote desktop into it from Vancouver
      * Software requirement will be a bit of a hiccup
      * Only 1 person who has administrator rights (Julie Hawkins)
        + Let anyone of us know what software we would like
        + Ask Julie to do instillations
      * Not sure about administrative privileges installing R packages – need to try
      * Windows machine – Martin to send instructions on how to remote in
      * **ACTION – tell martin what software we think we need**
* Office space
  + Office manager on vacation, still waiting
  + Martin contacted government real estate people. The office Nasim is in are the only office open to having people of other ministries in their space

# Schedules/Vacation time – 5 mins

* Nasim has no scheduled leave at this point
* Martin in Ottawa for work next week, be 3 hours ahead and less available
  + Martin will send other possible times to reschedule this meeting
  + May 16th & 17th and following week martin will be away on leave (back to work 27th )

# Feedback on Presentation – 15 mins

* Nasim
  + Liked the presentation
  + Agree with research questions
  + Questions about trained model (more domain specific wording, explain for non-tech people) – maybe have a visual to explain more for non-tech people
  + Initial model
    - Why SVM classifier
      * Disadvantage of bag of words
        + Doesn’t consider order of words
        + This is just the base model
      * More deep learner models for secondary models
    - Nasim’s project
      * Used logistic regression (one vs all)
* Martin
  + On the deliverables – relating to the 2nd deliverable
    - Would prefer a high-level summary or presentation compared to a lengthy analytical report

# Feedback on Research Questions – 10 mins

* Research questions (spoke about in previous section)

# Random Questions – 5 mins

* Do they have Qualitative data over time?
  + Slight variations have been asked over time (they do go back to 2007)
  + Assembling the data for this project, they only pulled the 2018 data
* Were the labels double checked by multiple people?
  + Inter coder reliability
    - Assign everyone the same comment to code and discuss how they got to the results. Some comments are more straightforward. Once they were comfortable together, they would code largely individually but being consistent
    - When they crossed less certain comments, they would consult each other
    - More informal, no inter-rater metrics calculated

# Upcoming Deadlines – 5 mins

* Proposal report will be sent on Friday by noon
  + This will include all the important milestones for the project