PREDICTIONS

Term1 Score  
Term2 Score  
Term3 Score  
Term4 Score  
Term5 Score

N-GRAM  
DATABASE

PREDICT

SHINYAPP

Input  
Phrase

CVS Files for  
Uni-Grams  
Bi-Grams,  
Tri-Grams  
Quadgrams &  
Quintgrams

GENERATE  
PHASE

PREDICT

PHASE

ORIGINAL DOCS  
Blogs Articles Tweets

**CREATE**  
Spare Document Feature Matrix

EXTRACT  
Top n-gram Terms

**TIDY CORPUS**  
Extract 5%  
Lowercase  
Tokenize  
Remove Non-Alpha  
Remove URL/Tags