In short:

I have used an ensemble technique which will combine results of three algorithms (Rake, TF IDF, AKE)

I found that shortcoming of a particular algorithm is compensated by other algorithms.

Longer version:

1. TF IDF:

   My approach for TF IDF:
   - Text Pre-processing: Removing stopwords, expanding contractions, performing lemming
   - Leverage Inverse document frequency to normalize term frequency.
   - Keywords decided based on score.

   Example output:

   https://www.reddit.com/r/news/comments/6bac2a/cable_lobby_conducts_survey_finds_that_americans/

   neutrality , isps , government , internet , companies , comcast , republican , regulations , isp , cruz

2. RAKE:

   RAKE – Rapid Automatic Keyword Extraction
   I have adopted the code for RAKE from https://github.com/aneesha/RAKE and modified it for my application accordingly.

   RAKE managed to produce good results which included phrases.

   Approach for RAKE:
   - Candidate selection
   - Properties calculation
   - Scoring and selecting keywords

   Example output:

   https://www.reddit.com/r/news/comments/6bac2a/cable_lobby_conducts_survey_finds_that_americans/

   free market , net neutrality , government regulation , obamacare internet , ted cruz , services , allowed , step , business , view

3. AKE – TextRank

I have adopted it from http://bdewilde.github.io/blog/2014/09/23/intro-to-automatic-keyphrase-extraction/

I have left this algorithm majorly unchanged. Currently, a lot of candidate keywords are repeated as in the example below. These duplicated may be removed to optimize performance.

This algorithm produced below satisfactory results. But shows potential if modified and tweaked properly.

Example output:

https://www.reddit.com/r/news/comments/6bac2a/cable_lobby_conducts_survey_finds_that_americans/

people , internet , net , internet net neutrality , net neutrality , neutrality , net neutrality regulations , internet companies , internet service , fcc net neutrality

4. Ensemble Method:

I have used the Weighted Majority Voting technique to obtain results.
Weighted Majority Voting technique is used while there are several algorithms running in parallel. It does not take into account the rank of the word from any algorithm

I have modified this algorithm to account for rank. I am using rank as weight to give more preference to the top most words in all three algorithms.

Approach:
- Gather the keywords from all algorithms (RAKE, AKE, TF IDF) sorted according to rank.
- Using rank as weight to calculate score of keywords
- Top 10 scores are the keywords

Ensemble technique produced excellent results including phrases and words.
Although, neutrality & net neutrality has been repeated in the below example, this method overcame shortcomings of single algorithms to obtain better results. It includes phrases and words, phrases are not repetitive and provides some insight to the document.

Example output:

https://www.reddit.com/r/news/comments/6bac2a/cable_lobby_conducts_survey_finds_that_americans/

internet , neutrality , net neutrality , isps , regulations , services , government , people , telecom companies , free market

Future Work:

Although ensemble technique gives insight into the link successfully, few words are still repetitive. This could be removed.
Since I am dealing with short text and ungrammatical sentences, I would take a look at semantic understanding from text.
Could explore other algorithms like LDA, variants of KEA ect.