International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 13th Apr 2022. Link :http://www.ijiemr.org/downloads.php?vol=Volume-11&issue=ISSUE-04

DOI: 10.48047/IJIEMR/V11/I04/19

Title **COVID-19 System Using Question Answering and Query-Focused Multi-Document Summarization**

Volume 11, Issue 04, Pages: 106-119

Paper Authors
R.M.Manohar Bhanu , T. Hemanth, M. Umesh Chandra, T. Priyanka

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COVID-19 System Using Question Answering and Query-Focused Multi-Document Summarization

*R.M.Manohar Bhanu , *T. Hemanth, *M. Umesh Chandra, **T. Priyanka
* B. Tech Student, Dept of CSE, Seshadri Rao Gudlavalleru Engineering Collage
**Assistant Professor, Dept of CSE, Seshadri Rao Gudlavalleru Engineering Collage
Manoharbhanu040@Gmail.Com,Hemanthtannidi.H@Gmail.Com,Umesh9999.Oe@Gmail.Com
Priyankacstau@Gmail.Com.

Abstract:

The Goal Of Caire-COVID Is To Accomplish The Latest Challenge Of Extracting The Various Scientific Papers Declared For COVID19 Via Addressing Prioritised Questions Based By The Society And Summarising Key Query Pertaining Info. It Chooses And Identifies Evidential Snippets From Existing Literature Based On A Question, Combining Extraction Of Information With QA And Question Focused Multi document Summarization Approaches. The Employment Of Query focused Abstractive And Extractive Multi-Document Summarising Algorithms Aids In The Provision Of More Important Facts Linked To The Inquiry.

Key Terms: Question And Answering, Multi document Summarizer, Abstractive Summary, Extractive Summary.

1. INTRODUCTION

A Large Number Of Research Studies Have Been Written And Made Freely Available To The Medical Community In Response To The COVID19 Outbreak. And There Are Rising Demands For Good COVID19 Information Management By Both The Medical And Scientific Community And The Public As A Result Of This Massive Number Of Research Papers. Strong Scientific Subjects Must Be Summarised So That The General Public May Easily Get Essential Background Information About The Coronavirus.

This COVID-19 Accessible Researcher Dataset, Which Comprises Over 160,500 Scientific Publications Regarding COVID19 And Pertaining Coronavirus, Provides A Chance For The NLP Community To React To The Demands. Furthermore, It Raises A Fresh Problem, Until It Is Hard To Retrieve Precise Info About Certain Scientific Topics And

Themes From A Huge Number Of Unlabelled Sources.

Caire-COVID Is A Multi-Document Summarization And Neurological Question Answering System, Was Developed To Satisfy The Demands And Challenges Of Academic Information Management Connected To COVID19. A Document Retriever Module Finds The Much More Files In CORD-19dataset Based On A User Query With Great Coverage Using QA Models, A Snippet Selector Module Then Shows The Answers Or Justifications For The Question, Provided The Appropriate Texts. Further, We Present A Question Focused Multi document Summarizer That Creates Abstractive And Extractive Summarizations From Several Recovered Answer-Related Paragraph Fragments Relevant To The Question by Fine-Tuning Them For QA And Summary, As Well As Our Own Adaptation Methods For The COVID-19.

Due To The Lack Of A QFS Dataset For COVID-19, The Summarizer Module's Performance Is Evaluated Using Two Query-Focused Summary Datasets, The DUC Datasets And The Debatepedia Dataset. The Much More Common Datasets For The QFS Job Are DUC Datasets, And Debatepedia Is The First Substantial Abstractive QFS Dataset. Previous Work Of The QFS Challenge Has Used A Query-Document Relevance Score Or A Query Attention Model To Include Query Relevanceand Concatenate Queries To Materials Into A Well Before Transformed framework, Or Use A Seq2seq Model. On Both Datasets, By Combining The QA Module's Solution Applicability Into The Summary Phase, The Query focused Multi document Summarizer Delivers Consistent ROUGE Score Growth Over Through The BART Based Baseline Approach On The Abstractive Job And The Leadto The Extractive Task.As A Result, We Expect That Our Suggested Summarizer Module Will Perform Well When It Comes To Query-Focused Summarising For COVID-19 Queries.

Extractive Techniques Attempt To Summarise Articles By Extracting Significant Lines Or Phrases From The Source Text And Stitching Together Parts To Generate A Condensed Version. The Retrieved Sentences Are Then Used To Construct The Summary.

The Technique Of Constructing A Brief And Succinct Summary Of A Source Text That Captures The Important Ideas Is Known As Abstractive Text Summarization. Additional Phrases And

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Sentences Not Found In The Original Text May Be Included In The Produced Summaries.

2. RESEARCH ON THE WORK

Multiple Methods Have Been Constructed To Aid Both Scholars And The General Public In Exploring Significant Material Linked To Covid19 Since The Covid-19 Accessible Research Dataset Was Provided By The Allen Institute For AI.CORD-19 Search Seems To Be A Search Function which Processes CORD-19 Dataset Using Amazon Comprehend Medical. On Top Of CORD-19 Dataset, Google Created The COVID19 Research Explorer, A Semantic Search Interface. Meanwhile, Covidex Usemulti-Stage Search Architectures To Extract Various Data Aspects.

The Wellai COVID-19 Research Tool, An NLP Medical Connection Engine, Can Build A Well-Organized List Of Medical Concepts using A Covid-19-Related Ranking Probability, And The Tm COVID Is A Tool For Extracting Bioconcepts From Covid-19 Literature And Summarising Them. It Offers Positive Pertinent Fragments & Summaries Depending Upon That Operator 'S Query. In Addition To Obtaining Data. The Website Also Provides Well-Structured And Succinct Information Regarding Covid-19.

3. ARCHITECTURE

The Caire-COVID System's Architecture, Which Is Made Up Of Three Primary Modules: Query-Focused Multi-Document Summarizer, Relevant Snippet Selector, And Document Retriever.

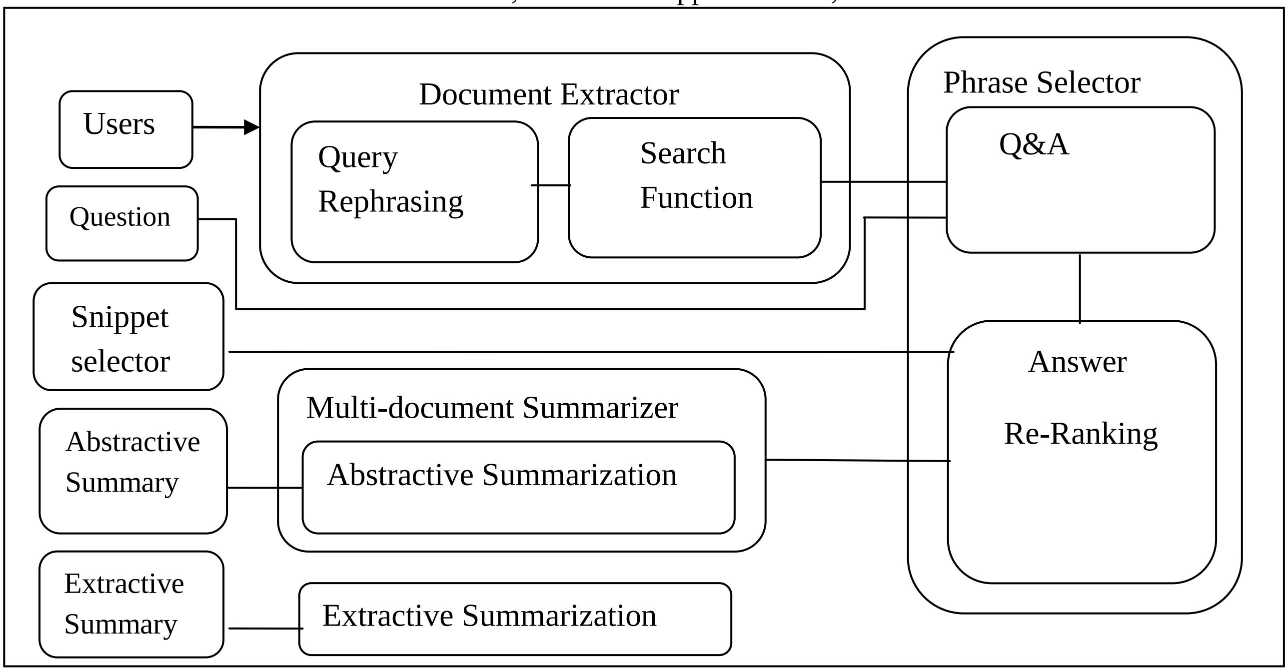

Figure 1: System Architecture

3.1 Document Extractor

The Matching Of A User Query Against A Set Of Free-Text Records Is Characterised As Document Retrieval. These Documents Could Be Anything With A Lot Of Unstructured Content, Such Newspaper Stories, Real Estate Records, Or Manual Paragraphs. User Queries Might Range In Length From Multi-Sentencedetailed Explanations Of Information Requirements To A Few Words.

Text Retrieval Is Sometimes Referred To As Document Retrieval, Or As A Subset Of It. Text Retrieval Is A Subset Of Information Retrieval In Which The Data Is Predominantly Stored As Text. Because Of The Personal Computer, Text Databases Become Decentralised. Text Retrieval Is An Important Field Of Research Nowadays Because It Is The Foundation Of All Internet Search Engines.

3.1.1Query Rephrasing

Although NLP Systems Can Understand Shorter Phrases More Easily, The Sub-Objective Module's Is To Deconstruct A User Question And Rewrite It Into A Series Of Shorter And Simpler Inquiries That Express The Same Idea. We Present Instances In Appendix while Dealing With Very Long And Tough Questions. This Feature Was Deleted From The Online System Since The Automated Methods Investigated Did Not Significantly Improve Our System's Performance.In The Future, More Automated Ways Will Be Investigated.

3.1.2 Search Function

The Search Tool Is A Bit Of Software That Uses Specialised Applications To Collect Data Such As Plain-Text, Page Layout, Meta Data Andother Designated Identifiers Of A Website's Content.

Anserini Is Used To Build The Search Function For Obtaining Basic Information database Containing Possible Documents. Anserini Is Just A Data Recovery Package Based On An Open Lucene Search Tool, Which Is Commonly Used To Create Sector Standard Search Function system. With The Help Of Lucene Indexing, Anserini Constructs A Simple Information Gathering Component. The Basic Algorithms For Ranking Such As Bag Of Words & BM25 Are Included To The Component. For Our Purposes, We Use Paragraph Indexing, In every Section For The Whole Text Of Every Article Throughout The CORD-19 Dataset, As Well As The Title And Abstract, Is Indexed Independently. To Any Of The

Questions, The Function Can Recall N Top Paragraphs That Match The Question.

3.2 Phrase Selector

A Phrase Is A Smallest Meaningful Body Of Text That Can Be Utilised To Determine The Document's Relevance Without Having To Open It. The Snippet Retrieval Work Builds On Previous Work In Focused Retrieval, Which Returns A Prioritised List Of Focused Components In Response To A User Query.

The Appropriate Phrase Selector Compiles A Set of Most Pertinent Solution Fragments given The Materials Provided, Emphasising Keywords That Are Essential. Given The Inquiries, Building A Cognitive QA Framework as A Assessor Of Factto Effectively Discover The Fragments Of Paragraphs Related To A Question. The Goal Of Quality Assurance Is To Predict Responses Or Evidence Text Spans Based On Relevant Paragraphs And Queries. The Solutions Are Underlined In The Articles, Which Are Then Re-Ranked Using A Well-Designed Score.

3.2.1 QA Using Evidence Selector

Selection Of Evidence Utilise A Combination Of Two QA Methods To Improve Generalisation And Domain Expertise The HLTC-MRQA Model And The Bio BERT Method.This HLTC-MRQA Framework Is An Xlnet-Based QA Machine That Was Trained On Six Distinct QA Datasets Using Number Of Co Learning. This Prevents Off To The Test Dataset And Enables For Non-Domain Data Generalisation, Resulting In Favourable Aspects. Appendix Contains Further Information. Rather Than Using COVID-19-Related Datasets To Perfect The QA Methodology, To Include The HLTC-MRQA Concept As A Scientific Proof Selector Into The System, Emphasis On Preserving The System's Generalisation Skills And Executing Zero-Shot QA. To Boost Efficiency, Integrate The HLTC-MRQA System With A Good Domain-Expert QA Framework.

Fusion As An Answer Apart Of Offering Short Responses, Offer Phrases As Outputs That Provide The Expected Answers To Improve The Readability Of The Responses. Both Pieces Of Evidence Are Maintained Once The Two Quality Assurance Methods Unearth Different Evidence From The Same Phrase. The Forecasts If The Two Methods Are Identical Or If There Is An Inclusion Link Between Them, They Will Be Mashed Together.

3.2.2 Re-Ranking answers

After Then, All Recovered Paragraphs Are Re-Ranked According To Their Significance To A Question. Score For Answer Confidence As The Answer's Score Of Assurance, Utilise The QA Systems' Prediction Probability. The Equation Computes Overall Aggregate Score Of Two QA Designs.

Score Based On Keywords To Determine The Matching Score Between A Question And The Returned Phrases, Employ Word Similarity. To Get This Score, We Use POS-Tagging To Pick Important Keywords From The Query, Only Words Containing The Tags NN, VB And JJ Are Considered. By Accumulating Of Phrase Counts With Total Number Of Particular Keywords things Are Mentioned In The Article, Two Matching Scores Can Be Calculated, That Are Labelled Sfreq And Snum, Respectively. Moderate Shorter Phrases By Obtaining A Sigmoid Value From The Phrase Length For The Wordmatching Score, As Well As Promote Phrases Having Higher Diverse Question Phrases. The Overall Perfectly Matched Score Is Then Computed.

Smatch =Λ1sfreq·Σ(L − Lc)+Λ2snum(I)Where L Is The Paragraph's Length And Lc Is A Length Constraint Because Of The Sigmoid Function's Influence, The Penalty Will Be Higher For Data Samples With Paragraph Lengths That Are Shorter Or Similar To Lc.

Re-Rank And Highlight: As Indicated In Equation, Using Both The Matching And Confidence Ratings, The Re-Ranking Score Is Calculated. The Relevant Excerpts Are Then Re-Ordered And Marked With The Respective Phrases.

Scorere−Rank = Smatch + Asconf (Ii)

3.2 Multi-Document summarization Based On Query

Multiple Document Summarization (MDS) Has Become A Non-Trivial Problem As The Amount Of Internet Information On The Issue Has Grown. By Providing A Succinct And Comprehensive Summary, The MDS Helps The User Understand A Vast Volume Of Information within A Brief Period Of Time. Furthermore, Query-Based Navigation By The User, The MDS System Generates A Standardised Summary, Which Includes The Most Important Information. The Traditional Summarising Techniques Concentrate On The Production Of Dynamic Query-Based Summaries.

The Query focused Multi document Summarizer Delivers Abstractive And Extractive Summaries Relevant To COVID19 Enquiries In Order To Accurately Deliver Relevant COVID19 Information.

3.2.1 Abstractive Text Summary

Fine-Tuning BART Our Abstractive Summarization Methodology Which Is Created On BART, Which Achieved Industry-Leading Results On Summarization Tasks Using The CNN/Daily Mail Datasets And XSUM. Because We Don't Have Any Additional COVID-19 Related Summarization Data. As The Fundamental Approach, Utilise The BART Framework Good Upon This CNN/Daily Mail Data.

Understanding Importance Of Relevance In Answers In Order To Offer Query-Focused Summaries, It Suggests Incorporating Response Relevance In Two Dimensions In The BART-Based Summarising Process. First, Rather Than Using The Document Retriever's Paragraphs List, As Input To The Multi-Document Summarizer, Use The Top K Paragraphs - Para1, Para2,..., Para K - Generated By The QA Function, These Are Re-Ranked Depending On The Significance Of Their Solution To The Question. The Anticipated Response Ranges From The QA Methods Are Then Mixed With Each Appropriate Phrase To Generate Answer Appropriateness, Rather Of Relying Just On The Re-Ranked Response Phrases To Create A Description. Attach The Question To The Conclusion Of The Insight As Well, Since This Has Been Shown To Function Successfully For The QFS Job.As A Result, The Summarization Model's Data Is C = Para1, Para2,..., Para K.

Summarization Of Multiple Documents As A Result, Each Para^I In C Might Originate From A Separate Post Which Emphasize In Various Perspective Of The Investigation. The Multi document Summary Is Created To Immediately Concatenating The Summary Of Every Para To Build An Ultimate Response Summary. There Would Be Some Redundancy, But We Feel It Is OK For Its Time Being.

3.2.2 Extractive Text Summary

To Create An Extraction Summary Based On A Question, First, Choose Response Phrases From A Variety Of Phrases As Possibilities, Which Make Up The QA Module's Response Spans. To Achieve Our Final Report, Re-Rank And Prefer The Best (K=3) Based On Their Response

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Relevance Score. A Response's Significance Score Is Calculated As Follows:

Sentence-By-Sentence Analysis To Get A Sentence Approximation, We Add The Contextual Annotations Recorded By ALBERT And Decrease By The Length Of The Sentence. Because Of A Combination Of Identity Levels And Intake Networks, This Version Can Capture Part Of The Lexical Definition Of The Text. Equation Determines The Representation H For A Sentence X = [W1, W2,..., Wn] With N Tokens, Equation Is Used To Calculate The Representation H.

$$e_{1:n} = ALBERT([w_1, w_2, \ldots.., w_n])$$

$$h = \frac{\sum_{i=1}^{n} e_i}{n} \qquad \text{(Iii)}$$

Similarity Calculationafter The Generation Of Phrase Representationword Vectors For Both The Answer Sentences And The Inquiry.The Cosine Identity Metric Is Utilised In This Study To Obtain The Consistency Score Amongst Them. Only The Three Largest Response Sentences For Each Inquiry Are Saved.

4. EXPERIMENTS

Let Us Run A Series Of Tests To Quantify The Efficiency Of Each Module And Illustrate The Use Of Our Approach.

4.1Question And Answering

We Run All Of The Trials In The QA Module With Hyper-Parameters Λ1 And Λ2 Set To 0.3, 11, Lc As 48, And A As 0.55.

4.1.1 Evaluation Of Quantity

To Boost Research Linked To COVID-19, This Dataset Evaluates The QA Module Performance On The Covidqa Dataset. The Covidqa Dataset Contains 125 Question-Article Pairings Connected To COVID19 For A One-Shot Evaluation Of The QA Mode's Transfer Ability.

Experiment With Different Settings Covidqa Dataset Evaluation Is Divided Into Two Parts: Textual Rank And Reliability Assessment. We Divided One Article Into N(N M) Paragraphs Since It Included Sentences. From Each Paragraph, One Statement Is Chosen As Proof For The Inquiry. In The Meanwhile, The Re-Ranking Scores For Each Sentence Are Determined. We Re-Rank The N Sentences Using The Re-Ranking Score After Evidence Selection. The QA Outcomes Are Evaluated Using MRR Accuracy Is Ranked First, And Retention Is Ranked Third.The MRR Is Calculated As Follows:

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \left\{ \frac{1}{rank_i}, 0 \right\}. \qquad \text{(Iv)}$$

With Ranked I Is The Place In The Initial Phrase At Which Perfect Answer Can Be Obtained Given One Article. If None Of The N Candidates Have A Golden Phrase, The Data Sample's Score Is Set To Zero.

Analysis Table Displays The Findings. Put The Models Through Their Paces With Basic Phrase And Key Queries. Overall Efficacy For Various Paradigms Varies Over Time, Displaying Their Opinions On Various Types Of Inquiries.

HLTC-MRQA System Containing Keyword Inquiries Outperforms The Model With Natural Language Questions In Terms Of Accuracy And Retention Percentages, Whereas The System Using Natural Language Inquiries Is Still Better Likely To Provide Appropriate Answers With A Similar Authority. The Biobert Model, On Either Side, Is Based On A Distinct Idea. As Two QA Methods Are Combined, Effectiveness In Regards Of Accuracy, Retention, And MRR Proportions Improves. Furthermore, The QA Module Surpasses T5 In Terms Of Accuracy Percent For Keyword Inquiries, And Our Approach Surpasses The Baseline In Terms Of Recall.

Question: What Are The COVID19 Risk Factors?

Answer: As Previously Stated, The Statistics Indicates Cardiovascular & Renal Illness, Obesity, Hypertension Are All Substantial Risk Factors Of COVID19 Issues." For Example, Having A Higher Adiposity Puts You At A Higher Chance Of Needing Breathing Assistance. In Reality, Obesity Is Associated With A Slew Of Risk Factors, Including Ectopic Fat Deposits, Poor Cardiopulmonary Kinematics, The Presence Of Various Illnesses, But Also Abnormal Pro inflammatory Reactions. Cognitive Problems, Especially Seizure, Are Considered To Become A Potential Element For Covid19, As Per The Centres For Infection Management And Prevention.

Extractive Summarization: Accurately Detecting Folks At High Potential Of Covid 19 Outcomes Might Aid Clinical Decisions And Global Safety Strategies, Preparatory Measures. In Diabetic Patients Hospitalised To Critical Care Units, COVID-19 Is Two To Three Times More Prevalent. High Blood Pressure Is A Significant Potential Element For Needing Breathing Help. Cognitive Disorders, Such As Seizure, May Be A Potential Component Towards COVID19, According To The Centres For Disease And Protection. At The Present, Neither

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Any Medical Record Of Seizure Has Been Recorded.

Abstractive Summarization: Despite The Dearth Of Data, The CDC Believes That Neurologic Disorders Like Seizures Might Be A Potential Element For Covid 19.Therefore, Viral Consequence, It's Unknown How Much The Research Group's Comorbidity Prevalence Differs From That Of SARS-Cov-2 Positive Persons Of The Same Age, And If These Illnesses Are Significant Hazard Elements Towards Serious Covid 19 Or Just Reflect Comorbidity Incidence In The General Population. What Hereditary And Non-Genetic Variables Impact Covid 19 Sensitivity In Different People.

Table 1: Example Question And Answering Pairings, As Well As The Abstractive And Extractive Summaries Produced By The System In Response To The CORD-19 Job Question

Template	Natural Language			Keyword Query		
	P1	R3	MRR	P1	R3	MRR
T5	0.280	0.400	0.410	0.213	0.373	0.365
Biobert	0.175	0.420	0.285	0.160	0.350	0.310
HLTC-MRQA	0.170	0.411	0.290	0.183	0.430	0.270
Ensemble	0.190	0.480	0.320	0.210	0.440	0.330

Table2: The QA Systems' Findings Were Obtained From The Covidqa Dataset. The T5 Model, Which Was Quite Well Utilizing The MS MARCO Dataset, Is The Strongest Effective Baseline.Nevertheless, Owing To Changes In Experimentation Parameters, The MRR Estimations Out Of These Simulations Are Incompatible With From Conventional Versions.

Template	R1			R2			RL		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
BART(C)	19.55	8.70	11.90	3.25	1.45	1.90	16.70	8.15	10.65
BART(C,Q)	20.45	9.23	12.40	3.60	1.59	2.25	17.49	8.60	11.20
BART(Q,C)	19.01	8.25	11.50	3.09	1.40	1.69	16.40	7.80	10.20
BART(A,Q)	20.10	8.90	12.10	3.40	1.39	2.01	17.30	8.31	10.90
BART(Q,A)	19.20	8.61	11.50	3.01	1.30	1.69	16.51	7.90	10.41
BART(C,A,Q)	21.95	10.09	13.40	4.25	1.90	2.51	19.10	9.41	12.20

Table 3: QFS Data Sourceoutcomes

Template	**DUC 2006**			**DUC 2007**		
	1	2	SU4	1	2	SU4
Lead	32.15	5.29	10.39	33.39	6.49	11.29
Extraction Process	34.50	6.49	11.19	35.50	7.80	12.10
BART(C_Nr)	35.80	6.30	11.41	37.90	8.09	12.99
BART(C)	38.00	7.65	12.80	40.69	9.29	14.39
BART(C,Q)	38.30	7.80	12.99	40.79	9.61	14.67
BART(C,A)	38.35	7.59	12.89	40.75	9.09	14.29
BART(C,A,Q)	38.30	7.69	12.90	40.50	9.25	14.40

Table 4: DUC Dataset Results

4.1.2 Examination of an Instance

Regardless Of The Fact That In The Majority Of Situations Where There Is An Acceptable Answer In The Paragraph, Two Versions Picked The Identical Line As The Ultimate Solution To A Problem, Indicating That They Had Dialectal Styles. The Biobert Approach Predicts A Selection For Any language Style, Whereas The MRQA Approach Predicts A Choice Towards A Particular Form Of Expression.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

4.2 Different Datasets

A File Containing One Or More Records Is Referred To As A Data Set. A Record Is The Most Fundamental Unit Of Information Utilised By A Z/OS Software. A Data Set Is A Named Collection Of Records.

The Structure Of DUC Datasets DUC 2006 And DUC 2007 Is Similar. To Meet The QA Model Input Criteria, We Divided The Documents Into Paragraphs Of 400 Words Or Less.

The Dataset Via Debatepedia The Information Comes From Debatepedia, A Collection Of Support And Opposing Arguments As Well As Remarks On Popular Discussion Subjects., And The Summaries Are One-Sentence Summaries Of Main Debate Issues. The Average Amount Of Words In Summaries, Papers, And Questions Is 11.20, 66.5, And 11 Respectively.

4.2.1 Outputs

For The Performance Comparison, Utilise ROUGE As The Assessment Metric.Tables 3 And 4 Illustrate The Outcomes With The Debatepedia QFS And DUC Datasets, Respectively. As Seen In The Two Tableson Both Datasets, Adding The Response Relevance Resulted In Constant ROUGE Score Boosts BART(C,A,Q) Over Most Other Options, Demonstrating The Usefulness Of Our Technique.In Addition, As Indicated In Table 4, Our Extractive Technique Consistently Improves ROUGE Scores Over The LEAD Baseline, And BART(C) Beats BART(C_Nr) In The Abstractive Scenario, Demonstrating That Re-Ranking Paragraphs Based On Their Response Relevance Can Assist Enhance Multi-Document Performance.

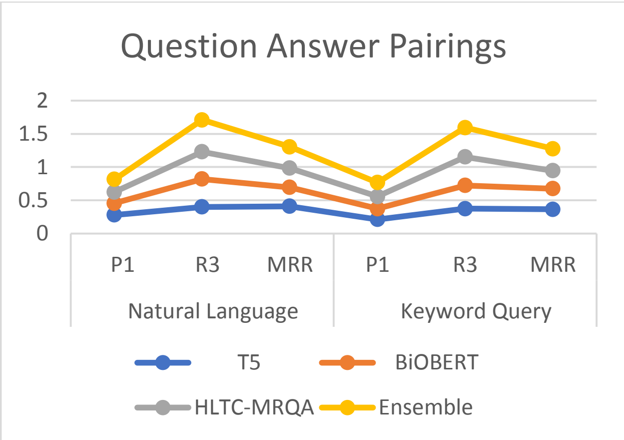

Figure 2: Q&A Parings

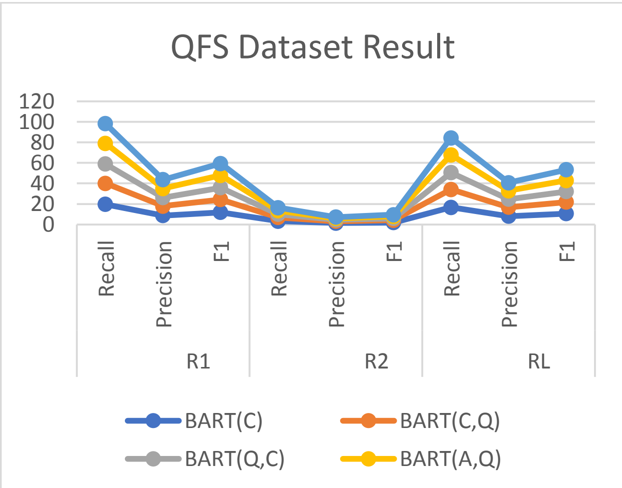

Figure 3: QFS Dataset Results

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

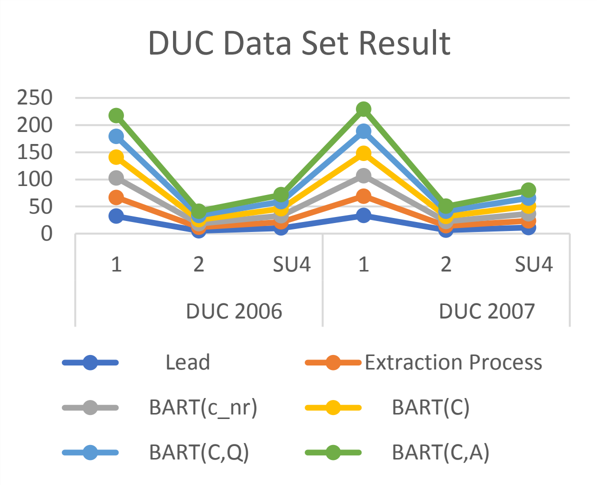

Figure 4: DUC Data Set Results

5. CONCLUSION

Caire-COVID Is A Multi-Document Summarising And Real-Time QA System That We Are Building For This Project.It Blends Data Extraction With Cutting-Edge Quality Assurance And Query-Driven Multi-Document Summarization Techniques. And It Provides Abstractive And Extractive Summaries For The Query.

References

[1] John Kinsman, John Angrén, Fredrik Elgh, Maria Furberg, Paola A Mosquera, Laura Otero-García, René Snacken, Tarik Derrough, Paloma Carrillo Santisteve, Massimo Ciotti, Et Al. 2018. Preparedness And Response Against Diseases With Epidemic Potenial In The European Union: A Qualitative Case Study Of Middle East Respiratory Syndromeand Poliomyelitis In Five Member States. BMC Health Services Research, 18(1):528.

[2] Naoto Kuroda. 2020. Epilepsy And Covid-19: Associations And Important Considerations. Epilepsy &Behavior, Page 107122

[3] Tom Kwiatkowski, Jennimariapalomaki, Olivia Redfield, Michael Collins, Ankur Parikh, Chris Alberti, Danielle Epstein, Illiapolosukhin, Matthew Kelcey, Jacob Devlin, Kenton Lee, Kristina N. Toutanova, Llion Jones, Ming-Wei Chang, Andrew Dai, Jakob Uszkoreit, Quoc Le, And Slav Petrov.2019. Natural Questions: A Benchmark For Question Answering Research. Transactions Of The Association Of Computational Linguistics.

[4] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, And Radu Soricut. 2019. Albert: A Lite Bert For Self-Supervised Learning Of Language Representations. Arxiv Preprint Arxiv:1909.11942.

[5] Md Tahmid Rahman Laskar, Enamul Hoque, And Jimmy Huang. 2020. Query Focused Abstractive Summarization Via Incorporating Query Relevance And Transfer Learning With Transformer Models. In Canadian Conference On Artificial Intelligence, Pages 342–348. Springer.

[6] Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, And Jaewoo Kang. 2020. Biobert: A Pre-Trained Biomedical Language Representation Model For

Biomedical Text Mining. Bioinformatics, 36(4):1234–1240.

[7] Mike Lewis, Yinhan Liu, Naman Goyal, Marjanghazvininejad, Abdelrahman Mohamed, Omer Levy, Vesstoyanov, And Luke Zettlemoyer. 2019. Bart: Denoising Sequence-To-Sequence Pre-Training For Natural Language Generation, Translation, And Comprehension. Arxiv Preprint Arxiv:1910.13461.

[8] Sewon Min, Victor Zhong, Luke Zettlemoyer, And Hannanehhajishirzi. 2019. Multi-Hop Reading Comprehension Through Question Decomposition And Rescoring. Arxiv Preprint Arxiv:1906.02

[9] Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman, And Phil Blunsom. 2015. Teaching Machines To Read And Comprehend. In Advances In Neural Information Processing Systems, Pages 1693–1701.

[10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, And Kristina Toutanova. 2019. Bert: Pre-Training Of Deep Bidirectional Transformers For Language Understanding. In Proceedings Of The 2019 Conference Of Thenorth American Chapter Of The Association For Computational Linguistics: Human Language Technologies, Volume 1 (Long And Short Papers), Pages 4171–4186.

[11] Deborah Cohen, Tai Huynh, Anne Sebold, Jean Harvey, Cory Neudorf, And Adalsteinn Brown. 2014. The Population Health Approach: A Qualitative Study Of Conceptual And Operational Definitions For Leaders In Canadian Healthcare. SAGE Open Medicine, 2:2050312114522618.

[12] Hoa Trang Dang. 2005. Overview Of Duc 2005. In Proceedings Of The Document Understanding Conference, Volume 2005, Pages 1–12

[13] Tal Baumel, Matan Eyal, And Michael Elhadad. 2018. Query Focused Abstractive Summarization: Incorporating Query Relevance, Multi-Document Coverage, And Summary Length Constraints Into Seq2seq Models. Arxiv Preprint Arxiv:1801.0

[14] Jonathan Berant, Vivek Srikumar, Pei-Chun Chen, Abby Vander Linden, Brittany Harding, Brad Huang, Peter Clark, And Christopher D Manning. 2014. Modeling Biological Processes For Reading Comprehension. In Proceedings Of The 2014 Conference On Empirical Methods In Natural Language Processing (EMNLP), Pages 1499–1510.