

TELUS SOW SUMMARIZATION

Yasamin abbaszadegan
Zonair Nadeem

Agenda

- 1- Problem Statement
- 2- Data SOW's Analysis
- 3- Summarization Techniques
 - Extractive
 - Abstractive
- 4- Results
- 5- Discussion

Problem Statement

In this project, you will be given several Statement of Work (SoWs) documents from TELUS which will be used for training/developing and testing purposes of this project. The SoW is the document that captures and defines all aspects of the project. It includes all the deliverables and the timetable for the project and is an extremely detailed document which lays the foundation for the project plan. This use case requires you to develop a NLP model to produce a summary of the SOW documents in roughly 200 to 400 words.

Corpus Preparation

Preprocessing of Corpus:

- Text cleaning.
- Total 14 Sections:
 - Description, Definition, Services, Term and Schedule, Place of performance and hours, Structure and Roles, General Responsibilities, Charges Expenses and Payment terms, Specific Service levels, Reports, Assumption and additional provisions, Addresses for Administration and invoicing, Agreement, Appendix 'A', Appendix 'B'.
- Extracted Sections based on best Results:
 - Services, Term and Schedule, Place of performance and hours, Structure and Roles, General Responsibilities, Charges Expenses and Payment terms, Assumption and additional provisions, Agreement.

Summarization Techniques: Used:

Extractive Summarization(TFIdf, TextRank, LSA):

The extractive approach involves picking up the most important phrases and lines from the documents. It then combines all the important lines to create the summary. So, in this case, every line and word of the summary actually belongs to the original document which is summarized.

Abstractive Summarization(T5 Small, Bart):

The abstractive approach involves summarization based on deep learning. So, it uses new phrases and terms, different from the actual document, keeping the points the same, just like how we actually summarize. So, it is much harder than the extractive approach.

TF-IDF Summarization

TF-IDF: A High weight in TF-IDF is reached by a high term frequency(in the given document) and a low document frequency of the term in the whole collection of documents.

Reference link:

<https://towardsdatascience.com/text-summarization-using-tf-idf-e64a0644ace3>

TF-IDF Summarization

- **TF = Term Frequency:**
 - $TF(t) = (\text{Number of times term } t \text{ appears in a sentence}) / (\text{Total number of terms in the sentence})$
- **IDF = Inverse Document Frequency:**
 - $IDF(t) = \log_e(\text{Total number of sentences} / \text{Number of sentences with term } t \text{ in it})$

Example:

Consider a sentence containing 10 words wherein the word **apple** appears once in it. The term frequency (i.e., TF) for **apple** is then $(1 / 10) = 0.1$.

Now, assume we have 1,000 sentences and the word *apple* appears in 10 of these. Then, the inverse document frequency (i.e., IDF) is calculated as $\log(1,000 / 10) = 2$. Thus, the TF-IDF weight is the product of these quantities: $0.1 * 2 = 0.2$.

We do this for all the words in the document, and then evaluate the importance of each sentence on sum of the TF-IDF score of the words in it.

TextRank Summarization

TextRank finds its roots associated with **Google's PageRank** (by Larry Page) used for ranking webpages for online search results. But before unfolding TextRank, we must understand PageRank & the intuition behind it:

- **PageRank:**
 - PageRank is used primarily for ranking web pages in online search results.
 - Suppose we have 4 web pages — w1, w2, w3, and w4. These pages contain links pointing to one another. Some pages might have no link – these are called dangling pages.
 - In order to rank these pages, we would have to compute a score called the **PageRank score**. This score is the probability of a user visiting that page.

Reference link:

<https://medium.com/data-science-in-your-pocket/text-summarization-using-textrank-in-nlp-4bce52c5b390>

webpage	links
w1	[w4, w2]
w2	[w3, w1]
w3	[]
w4	[w1]

PageRank(Cont')

The initialization of the probabilities is explained in the steps below:

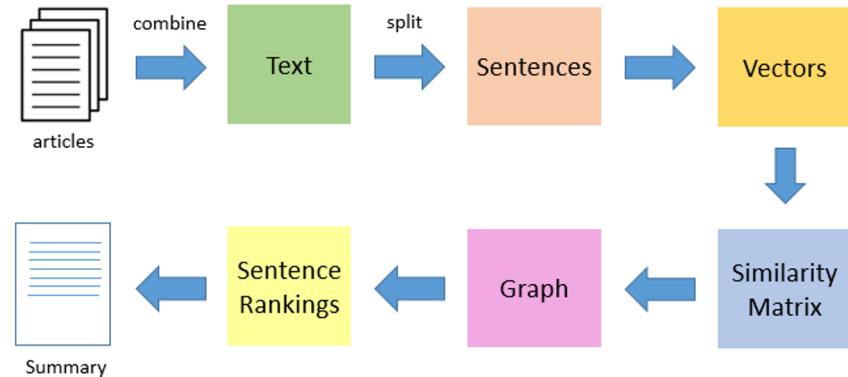
- Probability of going from page i to j, i.e., $M[i][j]$, is initialized with **1/(number of unique links in web page wi)**
- If there is no link between the page i and j, then the probability will be initialized with **0**
- If a user has landed on a dangling page, then it is assumed that he is equally likely to transition to any page. Hence, $M[i][j]$ will be initialized with **1/(number of web pages)**

Hence, in our case, the matrix M will be initialized as follows:

webpage	links
w1	[w4, w2]
w2	[w3, w1]
w3	[]
w4	[w1]

$$M = \begin{array}{l} \begin{matrix} & w1 & w2 & w3 & w4 \\ w1 & 0 & 0.5 & 0 & 0.5 \\ w2 & 0.5 & 0 & 0.5 & 0 \\ w3 & 0.25 & 0.25 & 0.25 & 0.25 \\ w4 & 1 & 0 & 0 & 0 \end{matrix} \end{array}$$

TextRank Summarization(con't)



- The first step would be to concatenate all the text contained in the articles
- Then split the text into individual sentences
- In the next step, we will find vector representation (word embeddings) for each and every sentence
- Similarities between sentence vectors are then calculated and stored in a matrix
- The similarity matrix is then converted into a graph, with sentences as vertices and similarity scores as edges, for sentence rank calculation
- Finally, a certain number of top-ranked sentences form the final summary

TextRank Summarization

Section by Section Summarization

Sample Document:

TI_SOW_58_2019_TM_MITS_Stratus_mock.docx

Overall total words from the sectioned document in the Summarized Version: 1002

Overall total words from the sectioned document before Summarization: 2452

Ratio to the sectioned document: % 40.864600326264274

Overall Original document words before Summarization: 4246

Ratio to the Original document: % 23.598681111634477

Servives
'Subject to the agreement the sowspecific scope of services shall include the following this sow provides a broad set of it services that are all delivered in a time and materials and staff augmentation delivery model.'

sumarized lenght of section 36
original lenght of section 122
Schedule

This statement of work shall commence on **january DATE** and shall end on **december DATE** with the period from sow start date to sow end date referred to as unless terminated earlier in accordance with the otherwise indicated to ti **days DATE** prior to the end of the initial or renewal term the extension of the end date and this sow shall be renewed for a period of **one month DATE** at the same terms and conditions as were in effect as of the date just prior to same notification process and telus option to nonrenewal for additional **one month DATE** terms shall apply at the end of each renewal limiting the terms of the agreement any time during the sow term telus may terminate this sow early for convenience by providing ti with a notice of **at least thirty calendar days TIME**.

sumarized lenght of section 137
original lenght of section 147
Place of Performance and Hours

Offshore **ORG** ti service representatives will perform services under this sow during various business and nonbusiness hours specific to each unique role.

sumarized lenght of section 21
original lenght of section 192
Role

'The ti manager will procure and manage ti service representatives as required in furtherance of its obligations under this sow and shall be responsible for providing qualified ti representatives with suitable personal development training education competence and skill to perform the services in a workmanlike manner.'

sumarized lenght of section 47
original lenght of section 375
Responsibilities

In the event of potential negative service impacts caused by telus the ti manager will make all reasonable efforts to work with the telus manager towards creating alternative riskmitigating solutions to deliver the services on time in scope and within the service charges contemplated addition without limitation ti shall also be overall responsible for the followingtelus responsibilities telus shall be responsible for the overall management and administration of the project and as such will retain overall responsibility and specifically the followingthe management of **third ORDINAL** party suppliers to the project except subcontractors to ti with reasonable advance notice providing ti with reasonable and timely access to telus employees and other resources facilities technical documentation and information systems in furtherance of its obligations under this sow provide telus to ti representatives provide appropriate access and login credentials to all necessary tools systems servers and other applicable resources if any **five days DATE** prior to initiation of services timely responses to key issues questions reviews and approvals as applicable to the extent required in the agreement providing appropriate access within the telus environment to appropriate and necessary engagement specific software tools and systems to enable ti to deliver its services under this sow notwithstanding the foregoing ti will not have access to **third ORDINAL** party tools systems servers facilities documentation and other such resources unless otherwise approved by telus left blank.

sumarized lenght of section 227
original lenght of section 383
Charge

The replacement ti representatives name id and other service relevant credentials must be sent by ti to the telus manager for interview if deemed required by the telus manager and for approval prior to such ti representative providing any services under this telus manager reserves the right to either participate in the interview process or delegate participation to another telus manager to act on behalf of the telus manager in the interview process ti will process invoices in line with the agreement and telus finance and ti finance agreed practices and proceduresfor ti representatives logging time through telus systems telus manager approval of time shall replace the **monthly DATE** invoicing approval productive hours ti will be expected to submit an invoice to telus **no later than five business days CARDINAL** after the end of each calendar **month DATE** once received telus will be required to review and approve or dispute the submitted billable **hours TIME** within **ten business days DATE** of receipt of the invoice costs will be recorded to telus cost centres via a **journal ORG** entry prepared by the ti finance team the **monthly DATE** invoice format will include program nameprogram descriptioncompany codecost centrebillable hoursproductive hoursovertime hourservice level credits service level bonuses eligible pre approved expenses.

sumarized lenght of section 201
original lenght of section 585
Assumptions

Ti acknowledges that it shall not store transfer transmit view access disclose process handle use or otherwise exploit directly or indirectly any restricted data outside of **canada GPE** nor will supplier provide any services involving the handling of restricted data from outside of restricted data means all personal information confidential information of telus as disclosing party that relates to any telus customer and telus customer data or information provided collected generated or otherwise known by supplier as a result of any actions under or in respect of this agreement and telus customer means any current or former customer of telus or of any telus data residency policy all of the system access and access to customer data for customers on do not touch list should be performed remotely at telus premises with no copying of mentioned data for these customers locally outside of **canada GPE**

sumarized lenght of section 143
original lenght of section 977
Agreement

'This sow and any change orders issued hereunder may be executed by the exchange of signed counterparts by facsimile transmission or electronically in pdf or similar secure format.'

sumarized lenght of section 28
original lenght of section 73

TF-IDF Summarization

Section by Section Summarization

Sample Document:

TI_SOW_58_2019_TM_MITS_Stratus_mock.docx

Overall total words from the sectioned document in the Summarized Version: 1547

Overall total words from the sectioned document before Summarization: 2673

Ratio to the sectioned document: % 57.87504676393566

Overall Original document words before Summarization: 4446

Ratio to the Original document: % 34.7953216374269

Servives
'Subject to the agreement the sowspecific scope of services shall include the following this sow provides a broad set of it services that are all delivered in a time and materials and staff augmentation delivery model.'

sumarized lenght of section 36
original lenght of section 122
Schedule

This statement of work shall commence on **january DATE** and shall end on **december DATE** with the period from sow start date to sow end date referred to as unless terminated earlier in accordance with the otherwise indicated to ti **days DATE** prior to the end of the initial or renewal term the extension of the end date and this sow shall be renewed for a period of **one month DATE** at the same terms and conditions as were in effect as of the date just prior to same notification process and telus option to nonrenewal for additional **one month DATE** terms shall apply at the end of each renewal limiting the terms of the agreement any time during the sow term telus may terminate this sow early for convenience by providing ti with a notice of **at least thirty calendar days TIME**.

sumarized lenght of section 137
original lenght of section 147
Place of Performance and Hours

Offshore **ORG** ti service representatives will perform services under this sow during various business and nonbusiness hours specific to each unique role.

sumarized lenght of section 21
original lenght of section 192
Role

'The ti manager will procure and manage ti service representatives as required in furtherance of its obligations under this sow and shall be responsible for providing qualified ti representatives with suitable personal development training education competence and skill to perform the services in a workmanlike manner.'

sumarized lenght of section 47
original lenght of section 375
Responsibilities

In the event of potential negative service impacts caused by telus the ti manager will make all reasonable efforts to work with the telus manager towards creating alternative riskmitigating solutions to deliver the services on time in scope and within the service charges contemplated addition without limitation ti shall also be overall responsible for the followingtelus responsibilities telus shall be responsible for the overall management and administration of the project and as such will retain overall responsibility and specifically the followingthe management of **third ORDINAL** party suppliers to the project except subcontractors to ti with reasonable advance notice providing ti with reasonable and timely access to telus employees and other resources facilities technical documentation and information systems in furtherance of its obligations under this sow provide telus to ti representatives provide appropriate access and login credentials to all necessary tools systems servers and other applicable resources if any **five days DATE** prior to initiation of services timely responses to key issues questions reviews and approvals as applicable to the extent required in the agreement providing appropriate access within the telus environment to appropriate and necessary engagement specific software tools and systems to enable ti to deliver its services under this sow notwithstanding the foregoing ti will not have access to **third ORDINAL** party tools systems servers facilities documentation and other such resources unless otherwise approved by telus left blank.

sumarized lenght of section 227
original lenght of section 383
Charge

The replacement ti representatives name id and other service relevant credentials must be sent by ti to the telus manager for interview if deemed required by the telus manager and for approval prior to such ti representative providing any services under this telus manager reserves the right to either participate in the interview process or delegate participation to another telus manager to act on behalf of the telus manager in the interview process ti will process invoices in line with the agreement and telus finance and ti finance agreed practices and proceduresfor ti representatives logging time through telus systems telus manager approval of time shall replace the **monthly DATE** invoicing approval productive hours ti will be expected to submit an invoice to telus **no later than five business days CARDINAL** after the end of each calendar **month DATE** once received telus will be required to review and approve or dispute the submitted billable **hours TIME** within **ten business days DATE** of receipt of the invoice costs will be recorded to telus cost centres via a **journal ORG** entry prepared by the ti finance team the **monthly DATE** invoice format will include program nameprogram descriptioncompany codecost centrebillable hoursproductive hoursovertime hourservice level credits service level bonuses eligible pre approved expenses.

sumarized lenght of section 201
original lenght of section 585
Assumptions

Ti acknowledges that it shall not store transfer transmit view access disclose process handle use or otherwise exploit directly or indirectly any restricted data outside of **canada GPE** nor will supplier provide any services involving the handling of restricted data from outside of restricted data means all personal information confidential information of telus as disclosing party that relates to any telus customer and telus customer data or information provided collected generated or otherwise known by supplier as a result of any actions under or in respect of this agreement and telus customer means any current or former customer of telus or of any telus data residency policy all of the system access and access to customer data for customers on do not touch list should be performed remotely at telus premises with no copying of mentioned data for these customers locally outside of **canada GPE**.

sumarized lenght of section 143
original lenght of section 577
Agreement

'This sow and any change orders issued hereunder may be executed by the exchange of signed counterparts by facsimile transmission or electronically in pdf or similar secure format.'

sumarized lenght of section 28
original lenght of section 71

TF-IDF: (Summary Results)

Original Schedule

this statement of work shall commence on january, 01st, 2019 and shall end on december, 31st, 2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.unless otherwise indicated to ti 30 days prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for a period of one month at the same terms and conditions as were in effect as of the date just prior to renewal. the same notification process and telus option to nonrenewal for additional one month terms shall apply at the end of each renewal term.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of at least thirty calendar days

TF-IDF Summary:

Schedule

This statement of work shall commence on january, 01st, 2019 DATE and shall end on december, 31st, 2019 DATE , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.unless otherwise indicated to ti 30 days DATE prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for a period of one month DATE at the same terms and conditions as were in effect as of the date just prior to renewal. without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of at least thirty CARDINAL calendar days.

summarized lenght of section 126
original lenght of section 149

TEXTRANK : (Summary Results)

Original Schedule

this statement of work shall commence on january, 01st, 2019 and shall end on december, 31st, 2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.unless otherwise indicated to ti 30 days prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for a period of one month at the same terms and conditions as were in effect as of the date just prior to renewal. the same notification process and telus option to nonrenewal for additional one month terms shall apply at the end of each renewal term.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of at least thirty calendar days

TEXTRANK Summary:

Schedule

This statement of work shall commence on january DATE and shall end on december DATE with the period from sow start date to sow end date referred to as unless terminated earlier in accordance with the otherwise indicated to ti days DATE prior to the end of the initial or renewal term the extension of the end date and this sow shall be renewed for a period of one month DATE at the same terms and conditions as were in effect as of the date just prior to same notification process and telus option to nonrenewal for additional one month DATE terms shall apply at the end of each renewal limiting the terms of the agreement any time during the sow term telus may terminate this sow early for convenience by providing ti with a notice of at least thirty calendar days TIME.

summarized lenght of section 137
original lenght of section 147

LSA Summary Results:

Original Schedule

this statement of work shall commence on january, 01st, 2019 and shall end on december, 31st, 2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.unless otherwise indicated to ti 30 days prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for a period of one month at the same terms and conditions as were in effect as of the date just prior to renewal. the same notification process and telus option to nonrenewal for additional one month terms shall apply at the end of each renewal term.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of at least thirty calendar days

LSA Summary:

Schedule

This statement of work shall commence on january DATE and shall end on december DATE with the period from sow start date to sow end date referred to as unless terminated earlier in accordance with the otherwise indicated to ti days DATE prior to the end of the initial or renewal term the extension of the end date and this sow shall be renewed for a period of one month DATE at the same terms and conditions as were in effect as of the date just prior to same notification process and telus option to nonrenewal for additional one month DATE terms shall apply at the end of each renewal limiting the terms of the agreement any time during the sow term telus may terminate this sow early for convenience by providing ti with a notice of at least thirty calendar days TIME.

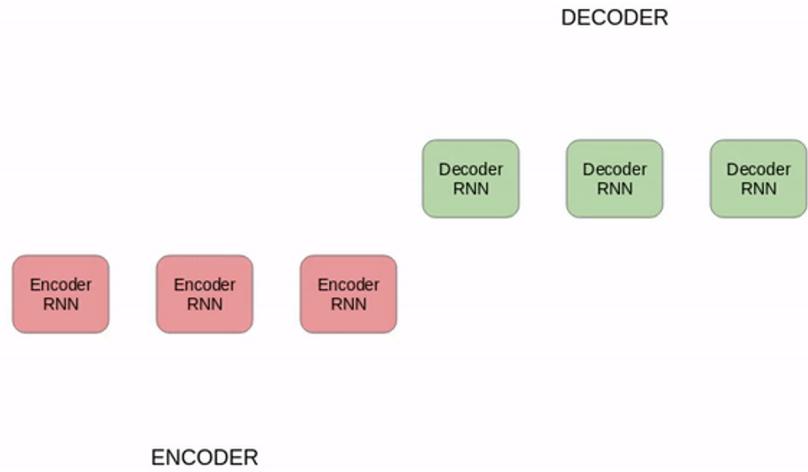
Abstractive Summarization Using BART and T5 Transformers

What are Sequence to sequence (seq2seq) Models ?

Sequence-to-sequence (seq2seq) models in NLP are used to convert sequences of Type A to sequences of Type B. For example, translation of German sentences to English sentences is a sequence-to-sequence task.

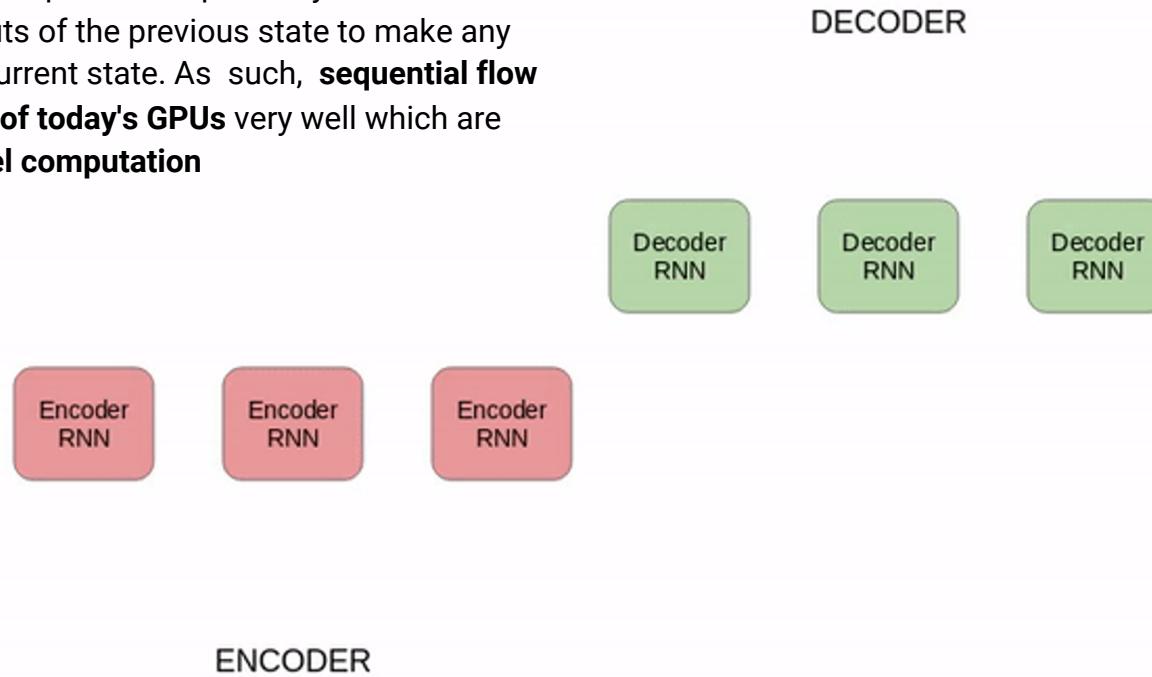
These sequence-to-sequence models are pretty versatile and they are used in a variety of NLP tasks, such as:

- Machine Translation
- **Text Summarization**
- Question-Answering System, and so on



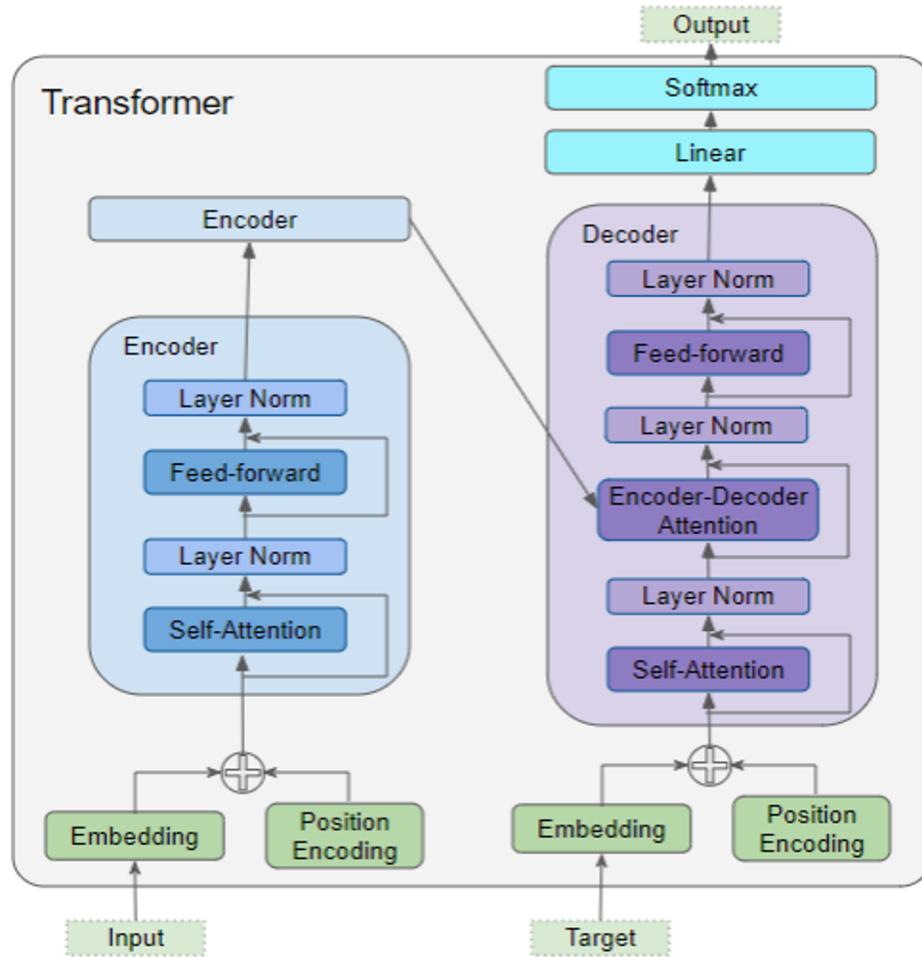
Sequence to Sequence Examples

For Seq2Seq Networks (**RNN(1982)**, **LSTM (1997)**) , the input data needs to be passed sequentially one after the other. We need inputs of the previous state to make any operations on the current state. As such, **sequential flow does not make use of today's GPUs** very well which are designed for **parallel computation**



Transformers

Introduced in 2017



Transformers

Steps for both Encoders and Decoders:

Step1(Input Embedding + Positional Encoding):

The Embedding layer encodes the meaning of the word.
The Position Encoding layer represents the position of the word.

Step2 (Multi self Attention):

Captures contextual relationships between words in the sentence

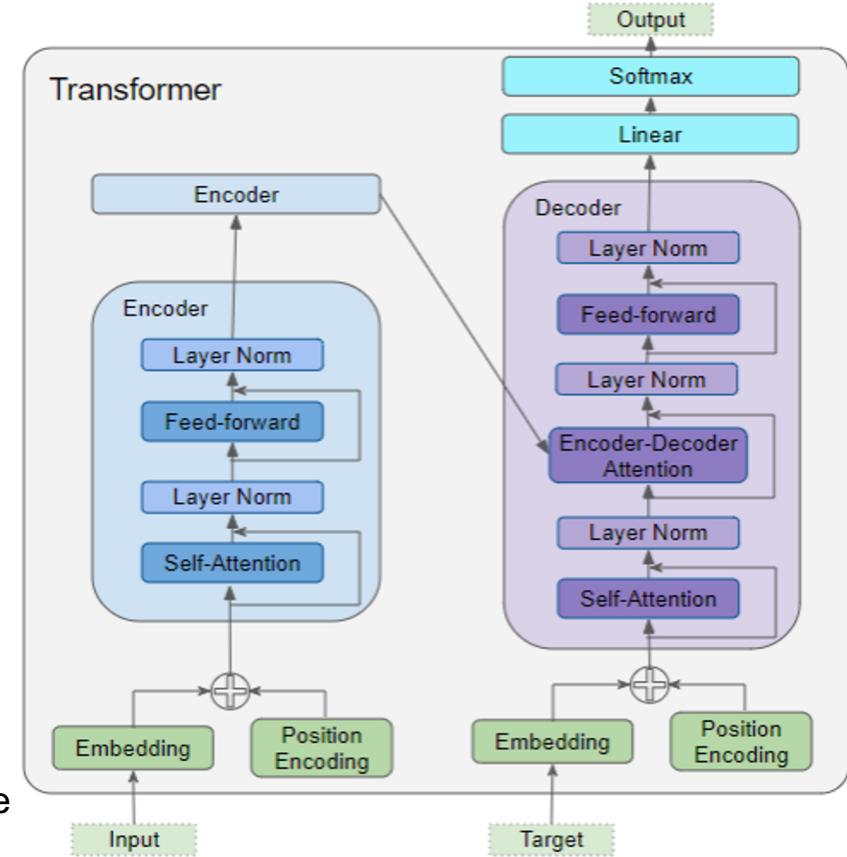
Step3 (Feed Forward):

Feed-forward nets are used in practice to transform the attention vectors into a form that is digestible by the next encoder block or decoder block

Next Step for Decoder:

Step4(Softmax Layer):

Softmax is used to create a probability vector which will help us determine the final output
(e.g. predicts the next word over multiple time steps until the end of sentence token is generated).



Step2 (Multi self Attention):



Dark colors represent higher attention (Image by Author)

Reference link:

<https://towardsdatascience.com/transformers-explained-visually-part-1-overview-of-functionality-95a6dd460452>

Transformers

Steps for both Encoders and Decoders:

Step1(Input Embedding + Positional Encoding): The Embedding layer encodes the meaning of the word. The Position Encoding layer represents the position of the word.

Step2 (Multi self Attention):

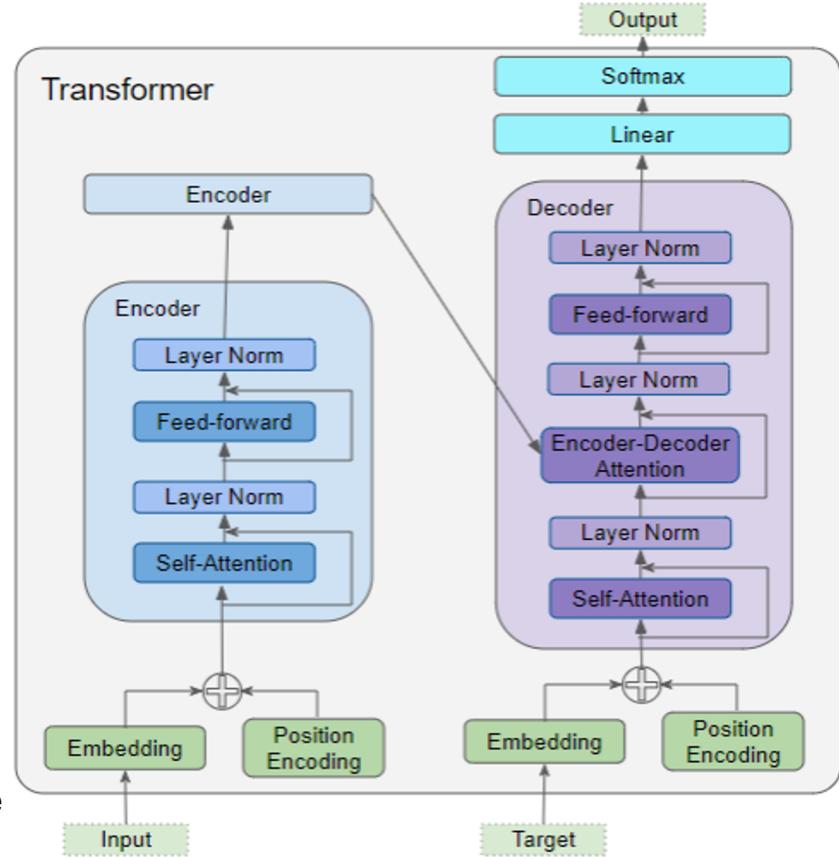
captures contextual relationships between words in the sentence

Step3 (Feed Forward):

feed-forward nets are used in practice to transform the attention vectors into a form that is digestible by the next encoder block or decoder block

Next Step for Decoder:

Step4(Softmax Layer): Softmax is used to create a probability vector which will help us determine the final output
(e.g. predicts the next word over multiple time steps until the end of sentence token is generated).



Parameters for Transformers

The decoder network generates the probability of occurrence of a word in the sequence. At each time step, the decoder has to make a decision as to what the next word would be in the sequence.

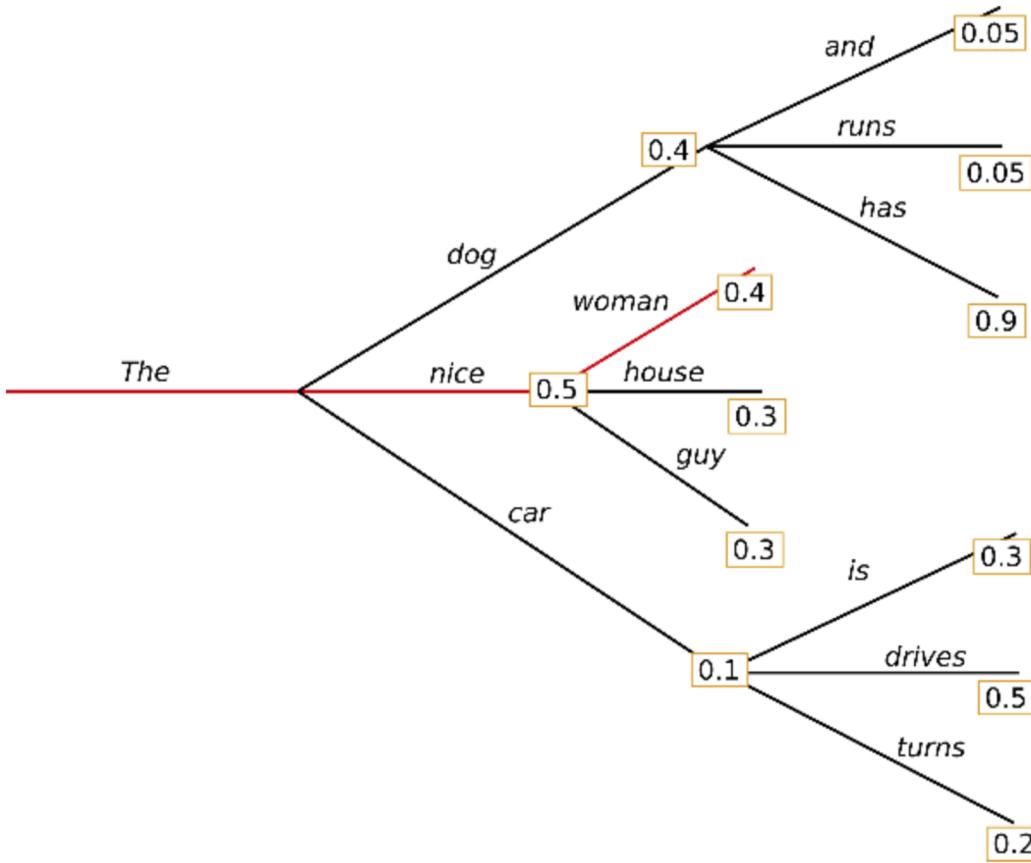
Transformer models have over 21 parameters, a few include:

**1-greedy search
2-beam search**

Reference link:

https://huggingface.co/transformers/main_classes/model.html

<https://huggingface.co/blog/how-to-generate>

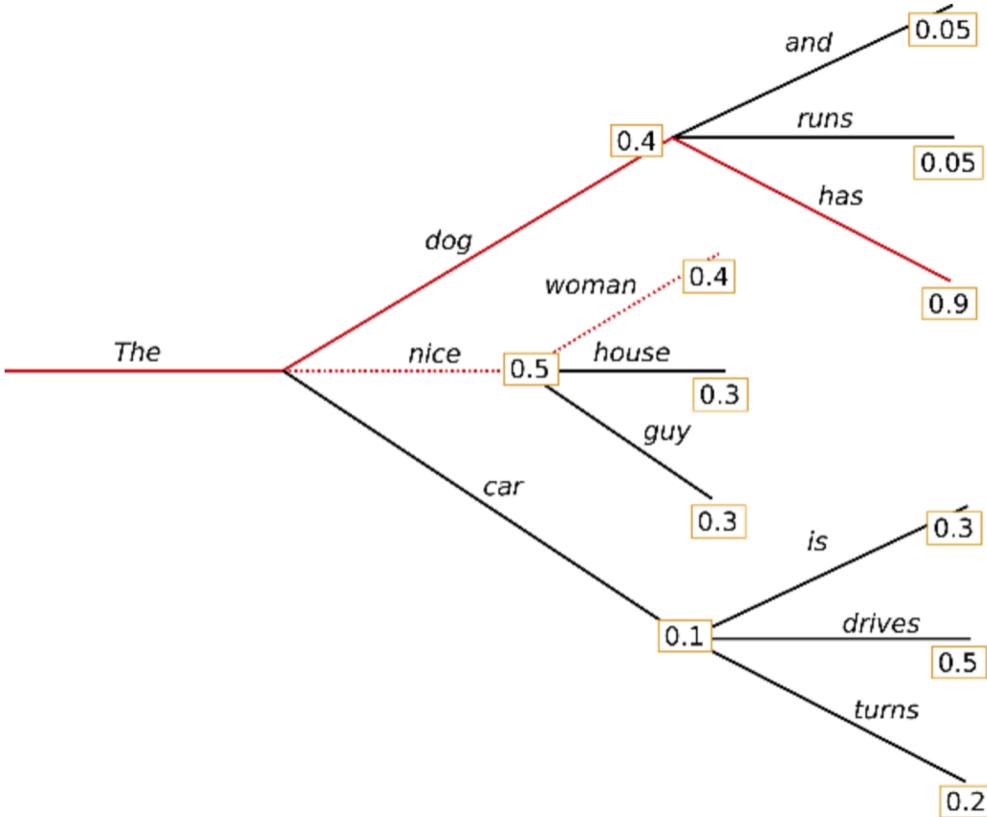


1-Greedy Search:

Greedy search simply selects the word with the highest probability as its next word

Reference link:

<https://huggingface.co/blog/how-to-generate>



2-Beam Search:

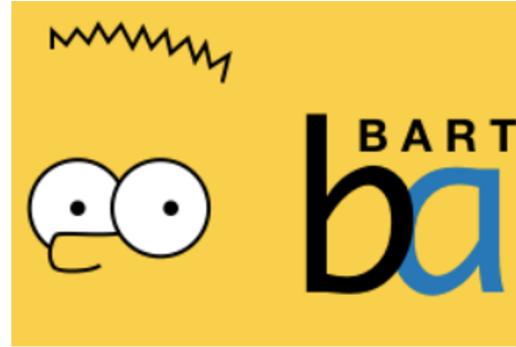
Beam search reduces the risk of missing hidden high probability word sequences by keeping the most likely num_beams of hypotheses at each time step and eventually choosing the hypothesis that has the overall highest probability.
Let's illustrate with num_beams=2:

Reference link:

<https://huggingface.co/blog/how-to-generate>

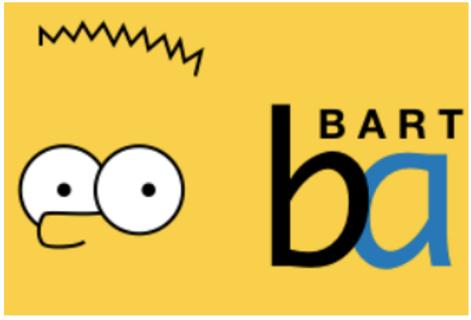
The **BART** model (**Bayesian Additive Regression Trees**) was presented in “*BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension*” by **Mike Lewis et al.**
2019

BART Transformers have many pre-trained models/weights, we chose:
“[facebook/bart-large-cnn](#)”



Reference link:

[https://huggingface.co/transformers/
model_doc/bart.html](https://huggingface.co/transformers/model_doc/bart.html)



Section by Section Summarization

Sample Document:

TI_SOW_58_2019_TM_MITS_Stratus_mock.docx

Overall total words from the sectioned document in the Summarized Version: 484

Overall total words from the sectioned document before Summarization: 2452

Ratio to the sectioned document: % 19.73898858075041

Overall Original document words before Summarization: 4246

Ratio to the Original document: % 11.398963730569948

Services

'subject to the agreement, the sowspecific scope of services shall include the following. this sow provides a broad set of it services that are all delivered in a time and materials and staff augmentation delivery model. the resources provided by ti service representatives are directed and managed by telus manager and their scope of duties is therefore open to change.'

summarized lenght of section 60
original lenght of section 122
Schedule

this statement of work shall commence on [january, 01st, 2019 DATE] and shall end on [december DATE], 31st,2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of [at least thirty

CARDINAL calendar days.

summarized lenght of section 72
original lenght of section 147
Place of Performance and Hours

ti shall perform the services at the following ti facilitiescanadian ti facilities none other [north american NORP] ti facilities none. offshore ti service representatives will perform services under this sow during various business and nonbusiness hours specific to each unique role. generally the work calendar will adhere to the telus [canada GPE] working calendar however the detailed schedule for each role and approval for adhoc nonworking days shall be set.

summarized lenght of section 68
original lenght of section 192
Role

'The parties shall appoint the following key personnel for the sow term. The ti manager will be responsible for the overall performance, delivery and management of services in respect of this sow. The telus manager shall manage and direct telus representatives and ti in accordance with this sow and the agreement. t.' []

summarized lenght of section 51
original lenght of section 375
Responsibilities

'ti responsibilities shall be responsible for the provision of all services in accordance with the service levels, if any, as attached to this sow per appendix 'a' , and as such, ti will retain overall project management responsibility for all ti service levels and ti service level impacting activities. ti will follow direction of the telus manager and other managers as from time to time designated by the telus manager.'

summarized lenght of section 69
original lenght of section 383
Charge

payment terms are set forth in the agreement, [article 8 based on section 1 LAW] , description, ti agrees to perform the services under this sow at the hourly rates quoted in [canadian NORP] dollars in section .the maximum total estimated fees are as follows cad [700.36 CARDINAL] . overtime for each calendar month, ti may charge telus for the daily hours of ti representatives overtime [hours TIME] .

summarized lenght of section 61
original lenght of section 585
Assumptions

ti acknowledges that it shall not store, transfer, transmit, view, access, disclose, process, handle, use or otherwise exploit, directly or indirectly, any restricted data outside of [canada GPE] . "restricted data" means all personal information: confidential information of telus, as disclosing party, that relates to any telus customer.

summarized lenght of section 46
original lenght of section 577

Agreement

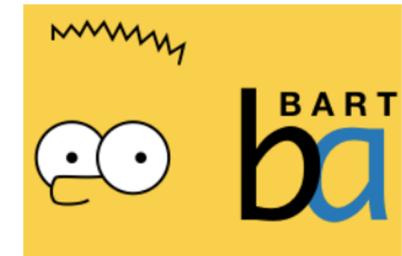
'the parties acknowledge and agree that the terms and conditions of the agreement shall govern this statement of work. this sow and any change orders issued hereunder may be executed by the exchange of signed counterparts by facsimile transmission or electronically in pdf or similar secure format. when taken together will constitute one and the same document.'

summarized lenght of section 57
original lenght of section 71

Overall total words from the sectioned document in the Summarized Version: 484
Overall total words from the sectioned document before Summarization: 2452
Ratio to the sectioned document: % 19.73898858075041
Overall Original document words before Summarization: 4246
Ratio to the Original document: % 11.398963730569948

Original Schedule

this statement of work shall commence on january, 01st, 2019 and shall end on december, 31st, 2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.unless otherwise indicated to ti 30 days prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for a period of one month at the same terms and conditions as were in effect as of the date just prior to renewal. the same notification process and telus option to nonrenewal for additional one month terms shall apply at the end of each renewal term.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of at least thirty calendar days

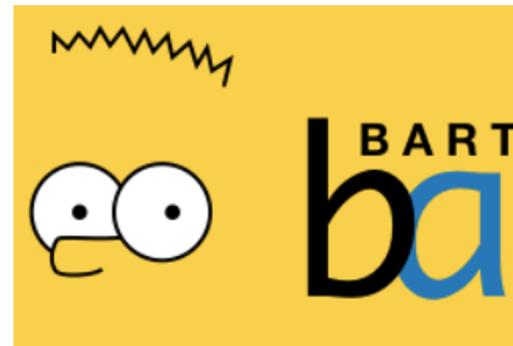


BART_Generic parameters

Schedule

this statement of work shall commence on **january, 01st, 2019 DATE** and shall end on **december DATE** , 31st,2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of **at least thirty CARDINAL calendar days.**

summarized lenght of section 72
original lenght of section 147



BART_Generic parameters

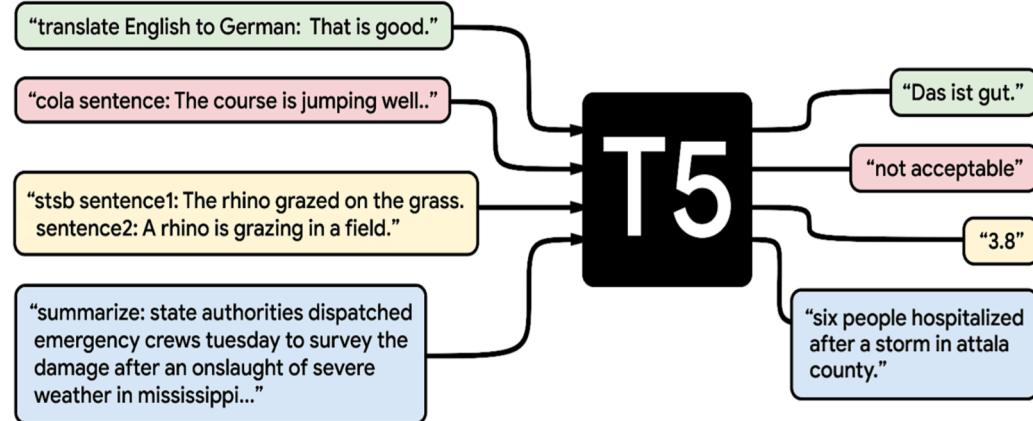
Charge

payment terms are set forth in the agreement, article 8.based on section 1 LAW , description, ti agrees to perform the services under this sow at the hourly rates quoted in canadian NORP dollars in section .the maximum total estimated fees are as follows cad 700,36 CARDINAL . overtime for each calendar month, ti may charge telus for the daily hours of ti representatives overtime hours TIME .

summarized lenght of section 61
original lenght of section 585

The **T5 model (Text-to-Text Transfer Transformer)** was presented in *"Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer"* by **Colin Raffel et al.** 2019

T5 Transformers have many pre-trained models/weights, we chose:
"T5-small" since it was computationally faster



Reference link: <https://ai.googleblog.com/2020/02/exploring-transfer-learning-with-t5.html>



Section by Section Summarization

Sample Document:

TI_SOW_58_2019_TM_MITS_Stratus_mock.docx

Overall total words from the sectioned document in the Summarized Version: 504

Overall total words from the sectioned document before Summarization: 2452

Ratio to the sectioned document: % 20.554649265905383

Overall Original document words before Summarization: 4246

Ratio to the Original document: % 11.869995289684407

Services

scope of duties is directed and managed by telus manager. scope of workjira issues bug fixes, enhancement management.testing of the workflows along with complete test cases writing; documentation of code and vro ORG workflowed documents - including tests case writing – are excluded from this sowspecific scope of services under this agreement nantermotto.

summarized lenght of section 52
original lenght of section 122
Schedule

this statement of work shall commence on january, 01st DATE . 2019 DATE and ends december 31th, 2019 DATE . the period from sow start date to end date will be renewed for one month DATE at the same terms and conditions as were in effect just prior to renewal unless otherwise indicated by telus 30 days DATE prior to the beginning or renew term if terminated earlier in accordance with the agreement - until the end of the initial or renewal term is terminated early – without limiting the terms of the agreement a time during

summarized lenght of section 90
original lenght of section 147
Place of Performance and Hours

offshore ti service representatives will perform services under this sow during various business and nonbusiness hours specific to each unique role. the work calendar will adhere to the telus canada GPE working calendar however the detailed schedule for each role and approval for adhoc nonworking days shall be set in consultation with the company manager.

summarized lenght of section 54
original lenght of section 192
Role

' the ti manager will procure and manage service representatives as required in furtherance of its obligations under this sow. telus manager shall cooperate with telus to perform reviews, ensuring the services are performed in a workmanlike manner – and serve as the single point of contact with respect to interfacing with the company.'

summarized lenght of section 53
original lenght of section 375
Responsibilities

' ti responsibilities will be responsible for the provision of all services in accordance with the service levels, if any, as attached to this sow. ti will follow direction of the telus manager and other managers as from time to time designated by the telus manager. the ti manager will proactively escalate issuesconcerns to the telus manager which may negatively impact ti's ability to satisfactorily provide the services.' ☐

summarized lenght of section 67
original lenght of section 383
Charge

ti agrees to perform the services under this sow at the hourly rates quoted in canadian NORP dollars in section 1 DATE . the maximum estimated fees for under this sow are as follows cad 700,361

CARDINAL shall provide the nonbillable ti service representatives in accordance as mutually agreed upon. ti may charge telus for the daily hours of ti representatives overtime hours TIME that are preauthorized and preapproved by the telus manager or authorized delegate.

summarized lenght of section 71
original lenght of section 585
Assumptions

ti acknowledges that it shall not store, transfer, transmit, view, access, disclose, process, handle, use or otherwise exploit, directly or indirectly, restricted data outside of canada GPE . telus customer data or information provided, collected, generated or otherwise known by supplier as a result of any actions under or in respect of this agreement. if applicable, any such equipment and software provided by ti must be configured, at its cost, for compatible use with telus systems, equipment, software, and network

summarized lenght of section 78
original lenght of section 577
Agreement

' the parties acknowledge and agree that the terms and conditions of the agreement shall govern this statement of work.counterparts. this sow and any change orders issued hereunder may be executed in counterparts, which when taken together will constitute one'

summarized lenght of section 39
original lenght of section 71

Overall total words from the sectioned document in the Summarized Version: 504

Overall total words from the sectioned document before Summarization: 2452

Ratio to the sectioned document: % 20.554649265905383

Overall Original document words before Summarization: 4246

Ratio to the Original document: % 11.869995289684407

Original Schedule

this statement of work shall commence on january, 01st, 2019 and shall end on december, 31st, 2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.unless otherwise indicated to ti 30 days prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for a period of one month at the same terms and conditions as were in effect as of the date just prior to renewal. the same notification process and telus option to nonrenewal for additional one month terms shall apply at the end of each renewal term.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of at least thirty calendar days

T5

T5_customized parameters

Schedule

this statement of work shall commence on **january, 01st DATE** . **2019 DATE** and ends **december 31th, 2019 DATE** . the period from sow start date to end date will be renewed for **one month DATE** at the same terms and conditions as were in effect just prior to renewal unless otherwise indicated by telus **30 days DATE** prior to the beginning or renew term if terminated earlier in accordance with the agreement - until the end of the initial or renewal term is terminated early – without limiting the terms of the agreement a time during

summarized lenght of section 90
original lenght of section 147

T5_Generic parameters

Schedule

this statement of work shall commence on **january, 01st, 2019 DATE** . the period from sow start date to sow end date referred to as. unless otherwise indicated to ti **30 days DATE** prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for **a period of one month DATE** .

summarized lenght of section 58
original lenght of section 147

T5

T5

T5_customized parameters

Schedule

this statement of work shall commence on **january, 01st DATE** . **2019 DATE** and ends **december 31th, 2019 DATE** . the period from sow start date to end date will be renewed for **one month DATE**

at the same terms and conditions as were in effect just prior to renewal unless otherwise indicated by telus **30 days DATE** prior to the beginning or renew term if terminated earlier in accordance with the agreement - until the end of the initial or renewal term is terminated early – without limiting the terms of the agreement a time during

summarized lenght of section 90
original lenght of section 147

T5_Generic parameters

Schedule

this statement of work shall commence on **january, 01st, 2019 DATE** . the period from sow start date to sow end date referred to as, unless otherwise indicated to ti **30 days DATE** prior to the end of the initial or renewal term, the extension of the end date and this sow shall be renewed for **a period of one month DATE** .

summarized lenght of section 58
original lenght of section 147

T5

BART_Generic parameters



Schedule

this statement of work shall commence on **january, 01st, 2019 DATE** and shall end on **december DATE** , 31st,2019 , with the period from sow start date to sow end date referred to as , unless terminated earlier in accordance with the agreement.without limiting the terms of the agreement, any time during the sow term, telus may terminate this sow early for convenience by providing ti with a notice of **at least thirty CARDINAL calendar days.**

summarized lenght of section 72
original lenght of section 147

T5_customized parameters

Charge

ti agrees to perform the services under this sow at the hourly rates quoted in **canadian NORP** dollars in **section 1 DATE**. the maximum estimated fees for under this sow are as follows cad **700,36t**

CARDINAL shall provide the nonbillable ti service representatives in accordance as mutually agreed upon. ti may charge telus for the daily hours of ti representatives overtime **hours TIME** that are

preauthorized and preapproved by the telus manager or authorized delegate.

summarized lenght of section 71
original lenght of section 585



T5_Generic parameters

Charge

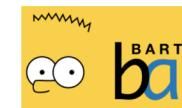
ti agrees to perform the services under this sow at the hourly rates quoted in **canadian NORP** dollars in **section 1 DATE**. all ratios will be reviewed as needed; changes can be made when both parties agree via the change management procedures as set forth in the agreement. ti may charge telus for the daily hours of ti representatives overtime **hours TIME** that are preauthorized and preapproved by the telus manager or authorized delegate.

summarized lenght of section 71
original lenght of section 585



BART_Generic parameters

Charge



payment terms are set forth in the agreement, **article 8.based on section 1 LAW**, description, ti agrees to perform the services under this sow at the hourly rates quoted in **canadian NORP** dollars in section .the maximum total estimated fees are as follows cad **700,36 CARDINAL**. overtime for each calendar month, ti may charge telus for the daily hours of ti representatives overtime **hours TIME**.

summarized lenght of section 61
original lenght of section 585

Challenges and Lessons Learned During Execution

Challenges:

1. Finding the optimal combination of content from the Project Sow document to use, to obtain meaningful insights for summarization. Going over all the documents and evaluating how are summarization models/Algo's capturing information from them.
2. What data cleaning steps to execute (Did not remove all punctuations)
3. Deciding which sections are most informative.
4. Understanding the property of each parameter for both Extractive and Abstractive models

Lessons Learned:

1. Garbage in will be Garbage out

Thank You

Any Questions??