WGU C951

Task 3

MACHINE LEARNING PROJECT PROPOSAL

Lucy Tran

001360516

March 14, 2022

Table of Contents

A. Project Overview

This proposal describes real-time language translation that supports two-way communication. The convenient programmed technology is ready on-the-go and available for any time usage in the form of an app. With compatibility for iOS and Android, simply download and utilize at your own convenience with the touch of your smartphone. Eliminating the complications of restricted language comprehensions essentially assists connections to thrive across world-wide.

A.1. Organizational Need

With so many languages in today's world, communication can prove to be difficult amongst those with differing speech. Offering a solution of how to effortlessly communicate with others by diminishing the language barrier with their own electronic mobile device that most already carry for distance communication. Suited for those in need of quick service, their efficiency is maximized and bear ease for any two interactions with different languages without human assistance.

A.2. Context and Background

Language barriers poses a communication problem for those less fluent. Over 7,000 languages are spoken around the world not to mention, the thousand types of dialects. It would be overly time consuming and impossible to manually learn them all. Since the early 1930s, machine translation has been improving where developers "teach" computers to translate text from one language to another without human input. Many types to machine translations have been created for different purposes. For example, Google is considered the leading translation engine that uses a process learning from repeated usage (Memsource, n.d.). As technology data collects and expands, artificial intelligence grows with knowledge and proficiency in turn delivers more accurate and to better satisfactory translation. Whether it's for traveling, conducting business, or on-the-go, reliability for language interpretation can be assured.

A.3. Outside Works Review

Machine translation evolution has sought to improve its accuracy levels significantly. As the fourth standing largest company in the world, Amazon came to life in 1994. Besides from being an online e-commerce, the company additionally offers Amazon Translate (AWS). According to their site, it is a neural machine translation service offering fast, high-quality, affordable, and customizable language translation. Benefits include accuracy and continuous improvement, easy integration, scalability, versatility, last but not least it is cost effective. As it continually improves, datasets expand to increase the validity; In turn, it permits a wider range of application and usage cases (Amazon Web Services, 2005). AWS has enabled multiple corporations to transform and expand with their translation services empowering efficient cross-lingual communication between users. Facilitating multilingual communication aids in preventing misunderstandings and misinformation.

App stores nowadays hold a variety of apps that offer translating services. Some are catered for more distinctive purposes while others are mainly for general use. Miles (2022) describes iTranslate Translator as being the best for viewing a translation dictionary, but that's not all its good for. It is the perfect tool for quick phrases, navigation in an emergency, or basic communication. The app has built-in phrasebook for swift look-up, not to mention voice-to-voice capabilities for over 100 languages. On the contrary, premium subscriptions and difficult interface compared to others are some listed cons for this tool. Offered free to download for both iOS and Android electronic devices, it's short-lived trial is limited to only a seven day period to which one would need to pay a subscription to continue the services.

With increasing competitively in the technological world, translator apps are not to be excluded. Minimizing the hurdle of language barriers, Hopwood (2022) reports:

> SayHi is dubbed as the "voice translator for everyone" and can be used for formal and informal speech. This is one of the most popular translation apps and it's free for iOS and Android. It has made the rounds in major traditional and online media such as the NBC Today show, TechCrunch, Lifehacker, and Gizmodo.

Moreover, this app is capable of translating text messages, voices, and conversational. Even more impressively, SayHi has a male or female voice setting with the ability to slow the speed of speech for better understanding. Furthermore, highlights mention supporting speech and text in 90 languages, plus an offline-mode. Enriched with learning on-the-go features, the app supports users of diverse background and intent.

A.4. Solution Summary

People need ways to understand each other without having to learn the entire language when interacting with those of a different linguistics. In today's world, the vast majority posses an electronic mobile device. Users may download the software that will allow them to connect with others effortlessly. Electronic devices having this app has the ability of automatic translation turns into a portable interpreter will save time and provide more confidence to those who want or need to converse with those of different lingual dexterity. Businesses and companies will benefit from machine language as a means to advance and prosper internationally.

A.5. Machine Learning Benefits

Machine translation continuously improves its capabilities as more data collects and technology advances over time. Speed is one of the key advantages. An average human needs at minimum of 30 minutes to translate 1,000 words (Brown, 2016). In comparison to a human, MT (machine translation) translates hundreds of thousands of words in a blink of an eye. Finding a translator to fit all your needs at your convenience in human form may prove quite hard at an incredible cost. Machine translation is advantageous as it offers speed and is cost effective compared to human translation in addition to its accessibility. The use of this artificial intelligence is applicable for a variety of usages to those of ranging ages. This will improve communication and tears down the language barrier that has limited people from being able to interact promptly.

B. Machine Learning Project Design

B.1. Scope

In Scope:

*** Enable users to access software via mobile**

*** Implementing a non-complex framework**

*** Outputs real-time accurate translation**

Out of Scope:

*** Logo design & product design**

*** Modify text to be translated more easily by translation system**

*** Improve accuracy & readability for users — *software updates***

B.2. Goals, Objectives, and Deliverables

Goals:

Construct a comprehensive mobile app-based software to promote communications globally between diverse speech that is user-friendly by implementing NLP with machine translation for effortless and uncomplicated user consumption.

Objectives:

* Visually appealing interface and easy navigation

* Mobile application platform

* Fast execution and response time of real-time translation

Deliverables:

* Simplistic and understandable language translating app

B.3. Standard Methodology

Development will follow the CRISP-DM methodology.

Business Understanding

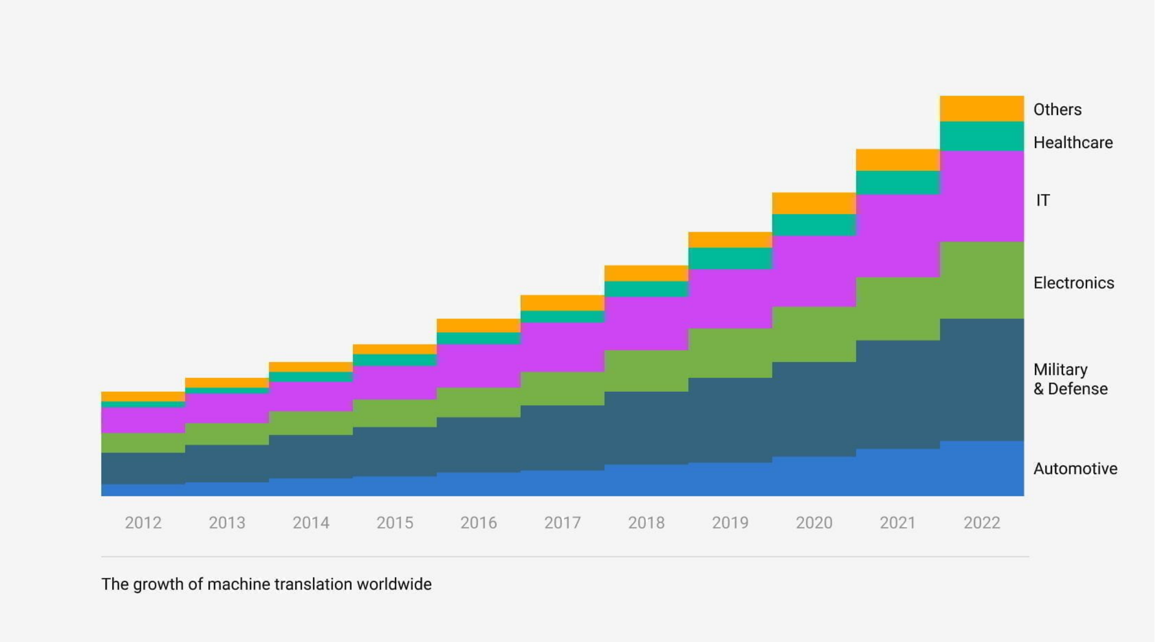
Design software to support language translation in efforts to improve accessibility and communication from mobile devices.

- Reduce need for human intervention for translation
- Usage of artificial intelligence with machine translation

Data Understanding and Preparation

Machine translation has increasingly grown and defines distinct advantages to its power. Translating purposes apply for casual conversing, business aspects, as well as academically. According to United Language Group (2018), MT improves turnaround times by approximately 35% or more depending on the scale of project. Not to mention, the tool is advantageous in decreasing the total cost of translation as it is effectively 100 times cheaper than human translation. With the ability to filter through extensive amounts of data, the performance and task is incomparable and impossible for any human to do in real-time or on large scales.

Figure 1 — Growth of machine translation worldwide.



Source: Becky Pearse (2020)

Advancements in development over-time has made its efforts to increase precision and accuracy. The evolution and structure of language processing systems has improved its consistency, producing newer types of machine translation suited and specialized for users.

Some types of Machine Translations:

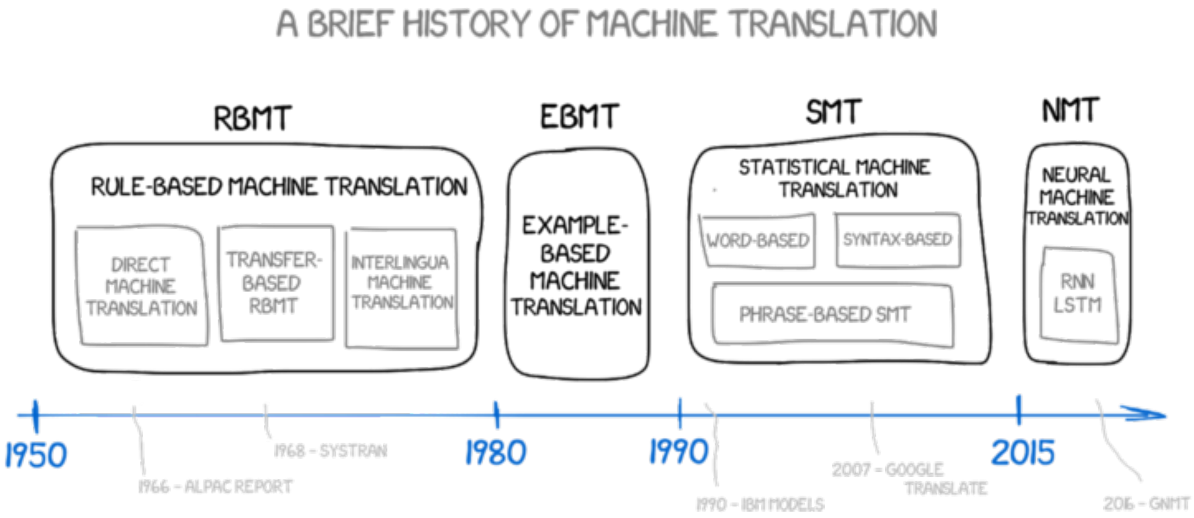
1. Rule-Based Machine Translation (RBMT)

2. Example-Based Machine Translation (EBMT)

3. Statistical Machine Translation (SMT)

4. Neural Machine Translation (NMT)

Figure 2 — History of Machine Translation


A BRIEF HISTORY OF MACHINE TRANSLATION
RBMT
RULE-BASED MACHINE TRANSLATION
DIRECT MACHINE TRANSLATION
TRANSFER-BASED RBMT
INTERLINGUA MACHINE TRANSLATION
EBMT
EXAMPLE-BASED MACHINE TRANSLATION
SMT
STATISTICAL MACHINE TRANSLATION
WORD-BASED
SYNTAX-BASED
PHRASE-BASED SMT
NMT
NEURAL MACHINE TRANSLATION
RNN LSTM
1950
1966 - ALPAC REPORT
1968 - SYSTRAN
1980
1990
1990 - IBM MODELS
2007 - GOOGLE TRANSLATE
2015
2016 - GNMT

Source: *Ilya Pestov (2018)*

Neural Machine Translation (NMT) is one of the newer translations. Told in *Google's Neural Machine Translation System: Bridging between Human and Machine Translation* (Wu et al, 2016)*:*

> The strength of NMT lies in its ability to learn directly, in an end-to-end fashion, the mapping from input text to associated output text. Its architecture typically consists of two recurrent neural networks (RNNs), one to consume the input text sequence and one to generate translated output text. NMT is often accompanied by an attention mechanism which helps it cope effectively with long input sequences. (p. 1)

NMT technology bases its learning on algorithms that constantly intake data and transforms with evolution. Vix (2019) explains, NMT has already impacted over 600,00 linguistics through over 21,000 translation service provision agencies. Using this intelligence for the creation of the software will improve speed in translation; Additionally, lowering overall costs. The machine translation selected for this project proves its effectiveness by continuously learning through data which in turn, does not limit users to distinct purposes such as business only.

Modeling

Encoding input to be translated into sequencing numbers that represent translated target sentences, NMT artificial intelligence uses a complex mathematical formula known as a neural network. Yip (2018) aids in the understanding of neural networks by thinking of the input as a signal with "information" on it to which it attempts to repeatedly refine and compress the information to match the desired output signal. The components to the framework includes an encoder, decoder. Studies completed help to better understand the algorithm for neural machine translation explains:

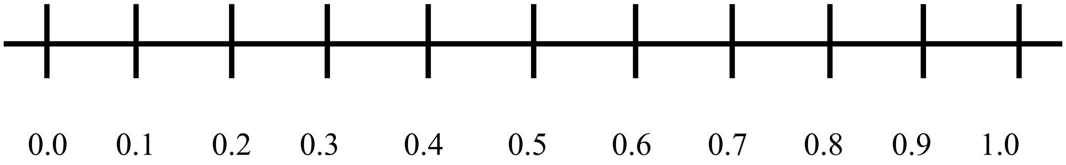
> A neural machine translation system is a neural network that directly models the conditional probability $p(y|x)$ of translating a source sentence, $x_1,...,x_n$, to a target sentence, $y_1,\ldots,y_m$. A basic form of NMT consists of two components: (a) an *encoder* computes a representation for each source sentence and (b) a *decoder* generates one target word at a time and hence decomposes the conditional probability as:
>
> $$\log p(y|x) = \sum\nolimits_{j=1}^{m} \log p\left(y_j | y_{<j}, \boldsymbol{s}\right)$$
>
> (Luong et al, 2015, p. 2)

When considering possible outcomes, NMT needs a search algorithm that will retrieve information quickly and efficiently to solve a given problem. For this project, the problem is finding the correct translation from one language to another. To determine whether that specific data exists within a structure, searching algorithms are used. If the data happens to exist, the algorithm will locate and retrieve the information (*Learn It: What are Searching Algorithms,* n.d.). By applying beam search algorithm, success in translation becomes more practical for the search space it must go through in order to quickly solve the translation.

Evaluation

In order to evaluate the model, a metric known as BLEU (BiLingual Evaluation Understudy) calculates a score to rate the machine-translated text based off its accuracy. The model quality score is a decimal between zero and one. The value of zero implies a low quality translation versus a value of one is high quality. The BLEU score however is not good at understanding meaning and grammar of sentences. For that reason, the score will only be used for single words and human evaluation will also be present with a score rating by a number ranging from one to three.

BLEU Scoring



[0] :: no overlap with reference

through

[1] :: perfect overlap with reference

Human Scoring

[1] :: zero meaning / nonsense translation

[2] :: some to most meaning retained, but may have grammatical mistakes

[3] :: completely consistent / perfect translation

Conclusion

Once the software has been created, a beta testing will release for user testing. Under one week of close observation, the product must maintain a high performance rate with low errors. With the beta testing, software engineers can ensure its satisfaction before the general public can fully access the product. Monitoring done by humans will be frequent in order for further improvements. Results from users usage are collected as well for reviewing and creating a report. Any low rated translations must be listed for urgent examination. A monthly audit will assess how well the software currently operates.

B.4. Projected Timeline

Timeline	Start Date	Due Date
Project Proposal	01/03/2023	01/07/2023
Technical Proof of Concept	01/08/2023	01/12/2023
Gather Requirements	01/13/2023	01/29/2023
Design	01/08/2023	02/18/2023
Develop	02/19/2023	09/27/2023
Analyze Results + Document	09/28/2023	10/11/2023
Submit for Review	10/12/2023	11/04/2023
Launch of Software		11/16/2023

Sprint Schedule

User Story	Task	Progress
As a developer, I want to utilize NMT in order to do translations in different languages.	Develop Design For Software	93%
	Collect Data to Train NMT Models	78%
	Evaluate Success of Interpretability	1%
As a developer, I have authorized access so that I can update/modify the software.	Release Beta Testing / Promo	0%
	Configurations + Monitoring	0%
	Document + Report Tests	0%
As a user, I can use this app on mobile device at my convenience.	App Store Available — iOS + Android	35%
	Install Contact/Help Page	4%

B.5. Resources and Costs

Resource	Description	Cost
Management	Project Manager Contact Management Business Intelligence	$15-$29/hr
Human Resources	Software Engineer Translators // SMEs (Literacy) Training + Development	$17-$55/hr
Hardware + Software Tools	Operating System CPU // Processor // RAM Intrusion Detection + DLP	$200-$4000
Workspace	Desks // Chairs Offices	$300-700
	Total Cost:	$100,000- $150,000

B.6. Evaluation Criteria

Objectives	Minimum Success Criteria
Ease of Use	80% Approval Rating From User
Annual Number of Downloads	200 Users
Response Time	approx. 100-300 milliseconds

C. Machine Learning Solution Design

C.1. Hypothesis

Offering users a simple, yet efficient and accurate translating app will result in removing the language barrier for communication in addition to achieving global clienteles and consumers when machine learning is utilize as the framework to create the translating software subsequently reducing the need for human interpreters.

C.2. Selected Algorithm

The algorithm selected is the beam search, commonly representing itself in a graph form. For input, this search has three ingredients to its recipe. This constitutes of the problem that needs solving, heuristic rules to for tidying up, and a memory space possessing bounded availability. In beam search, the superior states of each level is known as the beam width. The beam is where all the possibilities are stored. Once all nodes are explored, termination happens when the goal is fulfilled or could not be reached at all.

C.2.a Algorithm Justification

For NMT to locate the best translation, NMT will be used with a local search algorithm such as beam search algorithm. Widely coupled, beam search algorithm works heuristically, expanding all possible next steps and selects the most likely node, repeating until the end of the sequence. Beam Search Strategies for Neural Machine Translation (2017) expresses:

> In NMT, new sentences are translated by a simple beam search decoder that finds a translation that approximately maximizes the conditional probability of a trained NMT model. The beam search strategy generates the translation word by word from left-to-right while keeping a fixed number (beam) of active candidates at each time step. By increasing the beam size, the translation performance can increase at the expense of significantly reducing the decoder speed. (2.1.3)

Considering that this algorithm begins with the most likely matching words in the beginning of the sequence instead of random selection, performance enhances for the search process in comparison to other algorithms.

C.2.a.i. Algorithm Advantage

“The first use of a beam search was in the ***Harpy Speech Recognition System***, CMU 1976” (*Uses of Beam Search*, n.d.). The development of beam search focuses around optimizing best-fast search for the reason of lessening its memory demands. When beam width of the beam search algorithm expands, it requires less states to be eliminated for the search of solution. To avoid exceeding memory limit, memory held at the beam eradicates the most consuming node. For maximization of probabilities to find the targeted translation, this algorithm carefully makes selections of multiple alternatives to find the best match from input.

C.2.a.ii. Algorithm Limitation

Shortcomings of this algorithm includes its inability to guarantee optimality and may not achieve any results even after an extended amount of time. Successful outcomes aren’t always ensured, but having a larger beam width would secure more of an optimal path towards the desired target. Nonetheless, a larger beam width demands more memory and power for execution. Having a smaller beam width in return explores a lower quality of output. In events of inaccuracies, beam search influences the algorithm to skip over the shortest path to the objective.

C.3. Tools and Environment

Python is the primary programming language for the creation and execution. Software for this project will be compatible with iOS and Android operating systems. MySQL database will hold all the data. The application will be downloadable for users on their chosen electronic devices. Using open-source NMT toolkits on Github, this will help build the framework for the translation project. Integrating Google Translate API functions into our app and code will further aid in structuring text translations from one language into another. Open-source toolkits will also allow customization to its model that's built friendly for users.

C.4. Performance Measurement

Observing the performance of the algorithm, the worst-case space and time complexity of this algorithm is **O(B*m)**. This calculates how terrible it can do if the worst possible input was to run. **B** represents the beam width and **m** is for the maximum depth of any path within the search tree. Beam search is based off of best-first, however it only keeps *N* best items on queue that is prioritized. Starting at the very possible minimum of the beam width, it returns the first solution found. If there had been no solution, the beam then widens and repeats until attained. Performance is then calculated from the worst-case time complexity which demonstrates the time it took for the algorithm to execute its orders. Depending on the amount of memory used for execution, this computation is known as the space complexity.

D. Description of Data Sets

D.1. Data Source

The element of data sources contains techniques that provide valuable background and opportunities to improve and strategize this project. All information necessary to build the Android and iOS compatible app stem from an open-source database known as MySQL, powered by Oracle. This database is sufficient enough to handle the amount of users and continuous updates that is necessary for ML. Data is formatted for use by the algorithm to detect the correct output translation. MySQL centralizes the storage, access, and application of the data. Translation datasets taken from open-sources holds the collection of sentences and translations that have been pre-processed. Great Learning (2020) states: "Google has been the search engine giant, and they helped all the ML practitioners out there by doing what they are legends at, helping us find datasets. The search engine does a fabulous job at getting datasets related to the keywords from various sources, including government websites, Kaggle, and other open-source repositories". To add, Google also has filtering options for results. All of this will allow the collection of any datasets necessary for the languages presented in this app.

D.2. Data Collection Method

Data collection begins with Dataset Search by Google. Data collection techniques include searching new datasets and improving existing ones. The initial training of the machine learning model comes from selections of official source datasets where languages are translated to English that are free from the provider. Moving to improve existing data, ML model utilizes NMT methods therefore data can be generated to allow improved translation precision and speed.

D.2.a.i. Data Collection Method Advantage

The data selected are readily available and free for download. These models have been tested and pre-processed. Taking advantage of existing open-source expertises, this limits the time necessary to search for data. Guideline provisions for quality of some sources, especially ones from the government, are set to ensure high standards of material and data remain equal across the map.

D.2.a.ii. Data Collection Method Limitation

While these sources are free, they do not all prove to be always complete. Missing data arises as a problem in data collection as well as inaccuracies. Finding high-performing models for every single language out their is simply not possible. Data needs to be guaranteed 100% error-free for the algorithm, in essence draws a lot of time and knowledge that would require human assistance to verify. The need for further evaluations constitutes more tasks and costly resources.

D.3. Quality and Completeness of Data

Poor data is unacceptable for machine learning algorithms. Before the data is placed into the algorithm, it first needs to be prepped. Preparation beforehand for deployment will benefit in reduction of time and funds needed, along with assuring the data quality before deployment. Since data is downloaded from open sources, there is a need to format and convert the collection for overall consistency. As missing data values presents itself, the low level repetitive tasks are resolved with robotic process automation systems. It takes care of tedious tasks so that human handling can remain its focus on higher priority objectives.

D.4. Precautions for Sensitive Data

Securing data begins with setting policies and systems for appropriate handling. Data is classified based on sensitivity and access to the most sensitive data needs access restriction. Instilling the principle of least privilege restricts from random access. Limiting the access to need-to-know basis lowers the risk of threats to the data. The practice of Data Loss Prevention secures data from breaches and unwanted destruction as well as complying to regulations (*What is DLP*, 2022). Tools from DLP comprises of systems that alert attacker attempts, antivirus softwares, and firewalls. Security for data are instructed to all parties involved and regulations set forth violated will face ramifications.

References

3.1.3 searching algorithms. (n.d.). Retrieved May 15, 2022, from https://bournetocode.com/projects/GCSE_Computing_Fundamentals/pages/3-1-3-searc_alg.html

Amazon Web Services. (2005). *Amazon Translate*. AWS. Retrieved May 30, 2022, from https://aws.amazon.com/translate/

Brown, S. (2016, September 2). *How Fast Can a Translator Translate?* Blog. Retrieved May 16, 2022, from https://blog.languageline.com/how-fast-can-a-translator-translate-if-a-translator-translates-fast

Freitag, M., & Al-Onaizan, Y. (2017, June 14). *Beam Search Strategies for Neural Machine Translation*. Computation and Language. Retrieved May 15, 2022, from https://arxiv.org/abs/1702.01806

Group, U. L. (2018, April). *5 compelling reasons to use machine translation tools*. ULG'S LANGUAGE SOLUTIONS BLOG. Retrieved May 15, 2022, from https://www.unitedlanguagegroup.com/blog/compelling-reasons-to-use-machine-translation-tools

Hopwood, S. P. (2022, May 5). *Top 10 free language translation apps for Android and IOS*. Day Translations Blog. Retrieved June 2, 2022, from https://www.daytranslations.com/blog/top-10-free-language-translation-apps/

imperva. (2020, June 17). *What is DLP*. Data Loss Prevention (DLP). Retrieved May 15, 2022, from https://www.imperva.com/learn/data-security/data-loss-prevention-dlp/

javatpoint. (n.d.). *Uses of Beam Search*. Define Beam Search . Retrieved May 13, 2022, from https://www.javatpoint.com/define-beam-search

Learning, G. (2020, May 11). *Top 5 sources for analytics and machine learning datasets*. GreatLearning Blog . Retrieved May 26, 2022, from https://www.mygreatlearning.com/blog/sources-for-analytics-and-machine-learning-datasets/

Learn It: What are Searching Algorithms? 3.1.3 searching algorithms. (n.d.). Retrieved May 7, 2022, from https://bournetocode.com/projects/GCSE_Computing_Fundamentals/pages/3-1-3-searc_alg.html

Luong, M.-T., Pham, H., & Manning, C. D. (2015, September 20). *Effective approaches to attention-based neural machine translation*. arXiv.org. Retrieved May 15, 2022, from https://arxiv.org/pdf/1508.04025

Machine translation (MT): Everything you need to know. Memsource website. (n.d.). Retrieved May 15, 2022, from https://www.memsource.com/machine-translation/

Miles, B. (2022, January 3). *The 6 best translation apps of 2022*. Lifewire. Retrieved June 2, 2022, from https://www.lifewire.com/best-translation-apps-for-iphone-android-4774614

Pearse, B. (2022, April 25). *Human and machine translation: Both alive and kicking - and here to stay*. Translation & Localization Blog. Retrieved May 15, 2022, from https://www.smartcat.com/blog/human-and-machine-translation-both-alive-and-kicking-and-here-to-stay/

Pestov, I. (2018, March 1). A history of machine translation from the Cold War to deep learning. Retrieved May 15, 2022, from https://www.freecodecamp.org/news/a-history-of-machine-translation-from-the-cold-war-to-deep-learning-f1d335ce8b5/

Vix, B., About The Author Bill Vix More from this Author Bill Vix writes blogs, Bill Vix More from this Author Bill Vix writes blogs, Author, M. from this, Hanson, D., Parker, G., & McCutchen, M. (2019, January 23). *How close are we to getting machine translation perfected?* How Close are We To Getting Machine Translation Perfected? Retrieved May 15, 2022, from https://moneyinc.com/machine-translation/

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., … Dean, J. (2016, October 8). *Google's Neural Machine Translation System: Bridging the gap between human and machine translation*. arXiv.org. Retrieved May 15, 2022, from https://arxiv.org/pdf/1609.08144.pdf%20(7.pdf

Yip, S. (18AD). *What is Neural Machine Translation & How Does It Work?* What is Neural Machine Translation & How does it work? Retrieved May 15, 2022, from https://www.translatefx.com/blog/what-is-neural-machine-translation-engine-how-does-it-work?lang=en