

THE NETWORK

AUTHOR NAME: ASHINGTON WAWERU NJENGA

AFFILIATION: ASHFIN CONSULTANTS

Contents

INTRODUCTION 3

THE NETWAK STRATEGY 3

TECHNICAL BACKGROUND 3

STRUCTURE AND PURPOSE 3

KEY STAKEHOLDERS SOURCING OF NETWAK 3

NETWAK BUILD INFRASTRUCTURE 4

 Pretrained Model #1: Text-to-Text Transfer Transformer (T5) 6

 Pretrained Model #2: Neural Attentive Bag-of-Entities Model for Text Classification (NABoE) 7

 Pretrained Model #3: Rethinking Complex Neural Network Architectures for Document Classification
 9

NETWAK SOFTWARE DEVELOPMENT, SCALABILITY AND SUSTAINABILITY 10

1 TensorFlow 10

2 IBM Watson 11

3 Scikit-learn 12

4 Accord.NET 13

5 Torch 15

NETWAK ACTIVITIES	18
CONCLUSION.....	19
INDEX	19
References.....	21

INTRODUCTION

NETWAK is a global digital education strategy and network that aims to enhance the access of free education courses to the global disadvantaged individuals with an emphasis on gender sensitivity.

THE NETWAK STRATEGY

The main objective of the NETWAK strategy is to create an online authentication audit system that seeks and verifies needy students globally who are willing to learn and are hindered by lack of resources and lack of access to technology.

The strategy involves identification of sustainable funding sources for both technical manpower, computer and educational resources that make education networking and learning possible globally.

TECHNICAL BACKGROUND

The technical setup will require a server network to be established with protocols that enable authentication of user inputs collected from both computer and android phones on the NETWAK system. NETWAK is a non-profit making organization that provides free and low-cost certified education globally. NETWAK will be accessible on the internet to any android phone, tablet, PC and all laptop.

The network will be social media friendly to include networking capacity building on Facebook, Instagram, WhatsApp, Twitter, and many other social media platforms.

STRUCTURE AND PURPOSE

NETWAK provides linkages to education opportunities worldwide. Proper administrative privileges will be allocated to individuals chosen by the governing council under the leadership of the founder president. The president elects the ten members of the governing council and has the mandate to replace them at any time.

NETWAK is a global network that leverages the internet to provide authentication systems that vet the skills, qualifications, suitability and experience of network tutors and the credibility of individuals and communities intending to seek education via the network.

NETWAK creates linkages and networks with the aim of providing subsidized digital education opportunities by identifying scholarships for talents identified from needy early childhood to advanced college and university higher education students and tutors extending to industry professionals looking to further skills and training.

NETWAK collaborates with employer companies and industries globally to connect graduands with job opportunities and to inform curriculum development in relevant government institutions to meet industry employment standards.

KEY STAKEHOLDERS SOURCING OF NETWAK

NETWAK requirements target to seek and direct collaboration efforts sought with various global funding individuals and organizations who have vested interests in the education sector worldwide.

Some of the identified stakeholders include individuals and organizations like Richard Chandler, Bill and Melinda Gates foundation, the ELMA foundation, the Rockefeller foundation, Jeff Skoll, Rohini and Nandan Nilekani, Anzisha

Education Accelerator, Inside-Philanthropy, The Hundred Education Funders, The International Education Funders Group, edfunders, EWB Ventures, Firelight Foundation, Issroff Family Foundation, Mercy Corps Ventures, Acumen, African Visionary Fund, Allan Grey Orbis Foundation, Banyan Tree Foundation, Global Innovation Fund, IKEA Foundation, Kays Foundation, Lester Fund, LGT Venture Philanthropy, UBS Optimus Foundation, Wellspring Philanthropic Fund, Mastercard Foundation Edtech Fellowship, Co-Impact, Dubai Cares, Mastercard Foundation Young Africa Works, CDC Group, Big Win Philanthropy, DFID, Education Outcomes Fund for Africa & the Middle East, Global Partnership for Education, USAID, and the World Bank.

NETWAK BUILD INFRASTRUCTURE

NETWAK databases are built on virtual machines, mostly open access cloud server infrastructure build, and backed up in offline computer storage devices and data storage facilities comprising computer disks.

The NETWAK physical hardware topologies drive a communication strategy that enables information flow and access connected from an Inmarsat satellite.

Future prospects are in investment in GSM termination systems to ease costs of communications and create future income stream cashflows to enable financial sustainability of NETWAK.

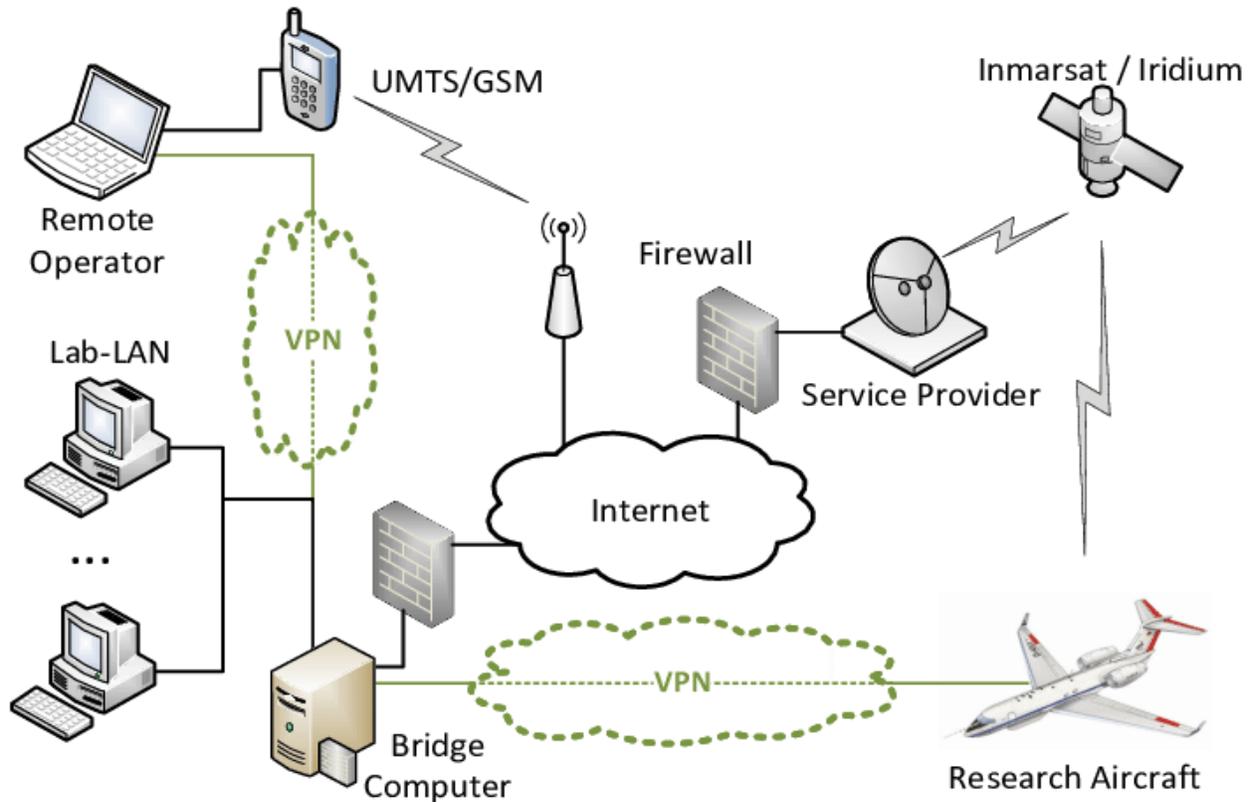


Figure 1 NETWAK PHYSICAL INFRASTRUCTURE TOPLOGIES

The NETWAK is built on highest level artificial intelligence specifications. Such specifications include a global URL website that runs on personal computers, android phones and tablets. A NETWAK prototype and application software are being developed on JavaScript platform.

System considerations include authentication processes embedded within the data collection, data processing, data capture and recording computerized global networks. NETWAK is built on cloud infrastructure with the capability to offer access and query to end users and administrators respectively. Query and access are enabled from any location or device worldwide (Vasilenko et al., 2020).

NETWAK offers capabilities for efficient and fast data collection and processing, and action response coupled with authentication against the artificial intelligence natural language processing programming layer which offers authentication against a collected database of similar word inputs.

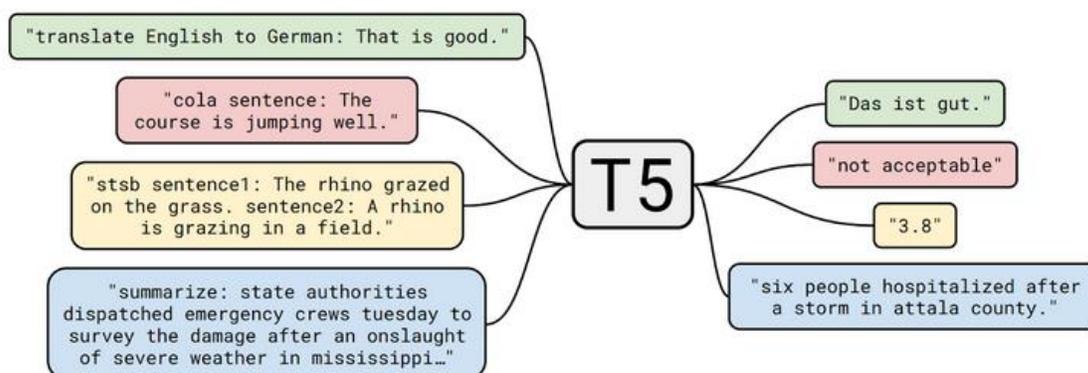
This allows for fast recognition of words that may allude to fraud in end user information collected and a vast possibility of continuous building-up or amassing of authentic database information for future cross referencing and authentication.

The natural language processing is built on a neural network with deep learning integration. This ensures the integrity of NETWAK in building capabilities in core requirements of the system including, building and compilation of an authentic database of similar word and phrase responses to recognize authentic pleas for access to education funding and counter fraudulent application for education help. The three neural networks pretrained models are discussed here (Goldberg, 2016).

Pretrained Model #1: Text-to-Text Transfer Transformer (T5)

I'll be honest – I had the most fun studying this model as compared to the others. Google's new Text-to-Text Transfer Transformer (T5) model uses [transfer learning](#) for a variety of NLP tasks.

The most interesting part is that it converts every problem to a text input – a text output model. So, even for a classification task, the input will be text, and the output will again be a word instead of a label. This boils down to a single model on all tasks. Not only this, the output for one task can be used as input for the next task.



The corpus uses an enhanced version of Common Crawls. This is basically scraped text from the web. The paper actually highlights the importance of cleaning the data, and clearly elucidates how this was done. Though the scraped data generates data of 20TB per month, most of this data is not suitable for NLP tasks.

Even after retaining only text content (pages containing markups, code content, etc have been removed), the corpus still has a size of a whopping 750GB which is much larger than most datasets.

Note: This has been released on TensorFlow too: [c4](#).

The task which is to be performed is encoded as a prefix along with the input. As you can see in the diagram above, be it a classification or a regression task, the T5 model still generates new text to get the output.

The T5 achieves SOTA on more than 20 established NLP tasks – this is rare, and taking a look at the metrics, it is as close to a human output as possible.

Task	Dataset	Metric	Metric Value	Global Rank
Question Answering	BoolQ	Accuracy	91.0	1
Document Summarization	CNN/Daily Mail	ROUGE-2	21.55	1
Linguistic Acceptability	CoLA	Accuracy	70.8	1
Semantic Textual Similarity	MRPC	F1	92.4	2
Sentiment Analysis	SST-2 Binary Classification	Accuracy	97.4	1

The T5 model follows up on the recent trend of training on unlabeled data and then fine-tuning this model on the labeled text. Understandably, this model is huge, but it would be interesting to see further research on scaling down such models for wider usage and distribution.

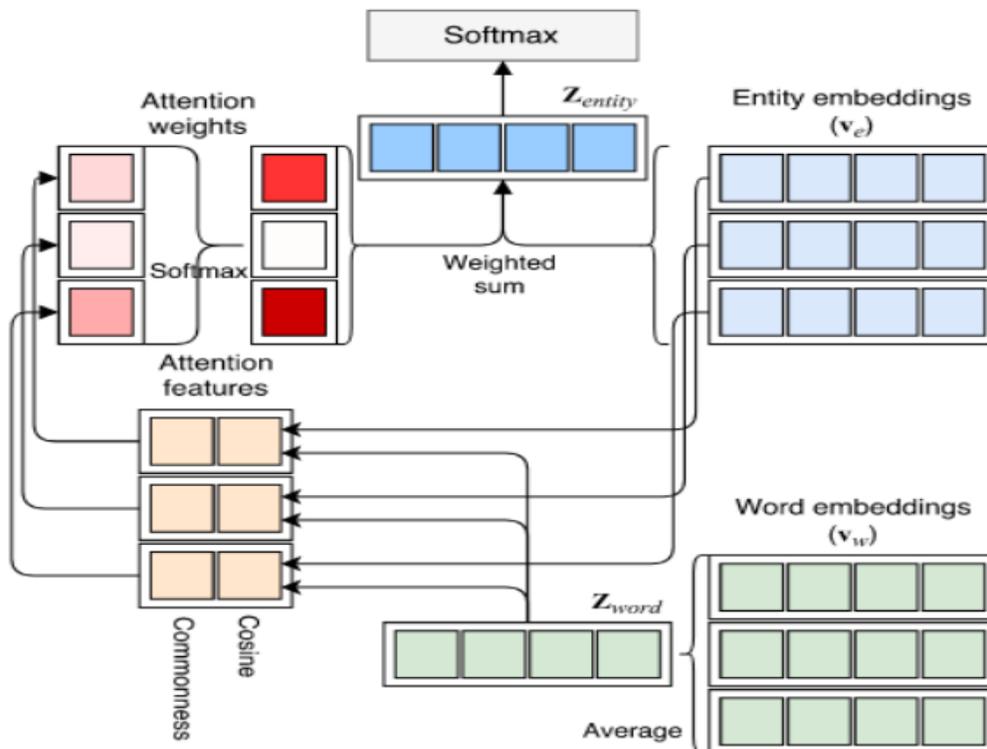
- Link to the Paper: [Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer](#)
- Github Link: <https://github.com/google-research/text-to-text-transfer-transformer>

Pretrained Model #2: Neural Attentive Bag-of-Entities Model for Text Classification (NABoE)

[Neural networks](#) have always been the most popular models for NLP tasks and they outperform the more traditional models. Additionally, replacing entities with words while building the knowledge base from the corpus has improved model learning.

This means that instead of building vocabulary from the words in a corpus, we build a bag of entities using Entity Linking. Though there has been research on this method of representing the corpus to the model, the NABoE model goes a step further by:

1. Using a Neural network to detect the entities
2. Using the attention mechanism to compute the weights on the detected entities (this decides the relevance of the entities for the document in question)



The Neural Attentive Bag of Entities model uses the Wikipedia corpus to detect the associated entities with a word. For example, the word “Apple” can refer to the fruit, the company, and other possible entities. Once all these entities are retrieved, the weight of each entity is calculated using the softmax-based attention function. This gives a smaller subset of entities which are relevant only to that particular document.

In the end, the final representation of the word is given by its vectorized embedding combined with the vectorized embedding of the relevant entities associated with the word.

The NABoE model performs particularly well on Text Classification tasks:

Task	Dataset	Metric	Metric Value	Global Rank
Text Classification	20NEWS	Accuracy	88.1	2
Text Classification	R8(SOTA)	Accuracy	97.9	1

- Link to the Paper: [Neural Attentive Bag-of-Entities Model for Text Classification](#)
- Github Link: https://github.com/wikipedia2vec/wikipedia2vec/tree/master/examples/text_classification

Pretrained Model #3: Rethinking Complex Neural Network Architectures for Document Classification

Now, it might appear counter-intuitive to study all these advanced pre-trained models and at the end, discuss a model that uses plain (relatively) old [Bidirectional LSTM](#) to achieve SOTA performance. But that was precisely why I decided to introduce it at the end.

Oftentimes, we miss the forest for the trees. We tend to forget that a simple well-tuned model might achieve just as good results as these complex deep learning models.

A combination of Bidirectional LSTM and Regularization is able to achieve SOTA performance on the IMDb document classification task and stands shoulder-to-shoulder with other bigwigs in this domain.

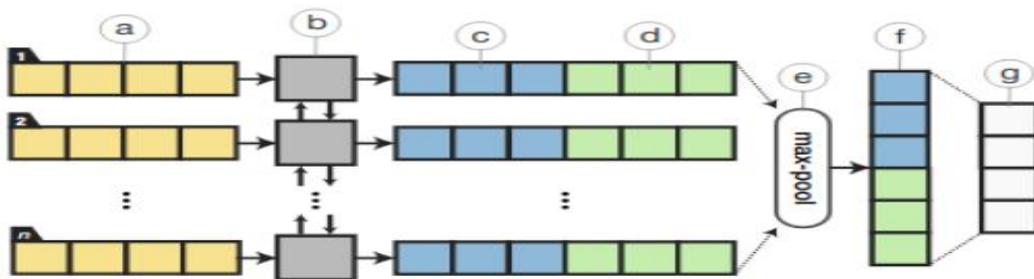


Figure 1: Illustration of the model architecture, where the labels are the following: (a) input word embeddings (b) BiLSTM (c, d) concatenated forward $h_{1:n}^f$ and backward $h_{1:n}^b$ hidden features (e) max-pooling over time (f) document feature vector (g) softmax or sigmoid output.

The most intriguing and noteworthy aspects of this paper are:

1. It does not use the attention mechanism
2. It is the first paper to use a combination of LSTM + regularization for document classification

This minimalistic model uses Adam optimizer, temporal averaging and dropouts to achieve this high score. The paper empirically compares these results with other deep learning models and demonstrates how this model is simple but effective and the results speak for themselves:

Task	Dataset	Metric	Metric Value	Global Rank
Document Classification	IMDb-M	Accuracy	52.8	1
Document Classification	Reuters-21578	F1	87.0	1
Document Classification	Yelp-5	Accuracy	68.7	1

This kind of model can be considered a novel approach for the industry where it is important to build production-ready models and yet achieve high scores on your metrics.

- Link to the paper: [Rethinking Complex Neural Network Architectures for Document Classification](#)
- Github Link: <https://github.com/castorini/hedwig>

NETWAK SOFTWARE DEVELOPMENT, SCALABILITY AND SUSTAINABILITY

The NETWAK scalability strategy incorporates considerations of freely accessible open-source software modelling under natural language and deep learning innovation. The free open-source resources to be used in future NETWAK software modelling are discussed in this section.

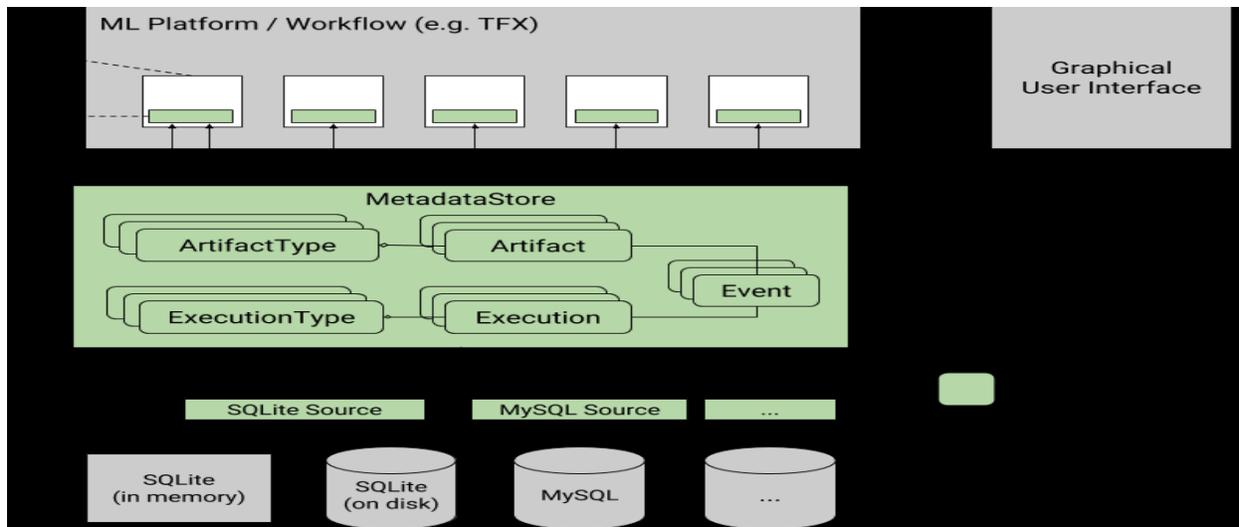
1 TensorFlow

TensorFlow is an open-source [artificial intelligence software](#) that helps you develop and train Machine learning models. It presents the library for high-performance numerical computation. Across a variety of platforms (CPUs, GPUs, TPUs) this free AI software allows easy deployment of computation due to its flexible architecture.

With this software, you can accomplish the power of data in your business by building advanced predictive modeling applications. This software makes use of data flow graphs to build models. In perception, understanding, prediction, creation, and classification, you can make use of this software.

You will find the primary use of this software in voice/sound recognition, text-based applications, image recognition, video detection, and time-series data. The time series algorithms of TensorFlow are mainly used in the field of finance, accounting, and enterprise /resource planning.

Several companies, like Dropbox, eBay, Twitter, Uber, and Intel, make use of this application. Behind several Google tools, this AI tool works as the engine. It contains Google photos and the speech recognition found in the Google app. This software works on various platforms like desktops, clusters, mobile, and edge devices, CPUs, GPUs, and TPUs due to its easy-to-use interface and architecture.



(Source: [TensorFlow](#))

Key features:

- Predictive Analytics
- Adaptive & Speech Recognition
- Data Ingestion
- Intuitive high-level APIs
- Easy model building
- Robust ML production
- Transfer Learning
- Multi-languages
- Workflow Automation

2 IBM Watson

IBM Watson is a free, open-source AI software that provides authority to the companies to speed up the research and discovery, calculate disruptions, and improve interactions. Several businesses are taking advantage of this software to study their data, gather intellectual property, insights, and predict their future performance easily. By using cloud-based IBM platform, organizations can make more informed decisions.

It is designed for enterprises from various sectors like IoT, consumer engagement, education, healthcare, energy, finance, work, and transportation. By using advanced AI and machine learning, enterprises can discover new opportunities and update themselves about possible regulations and risks. This app will allow you to increase your productivity.



(Source: [IBM Watson](#))

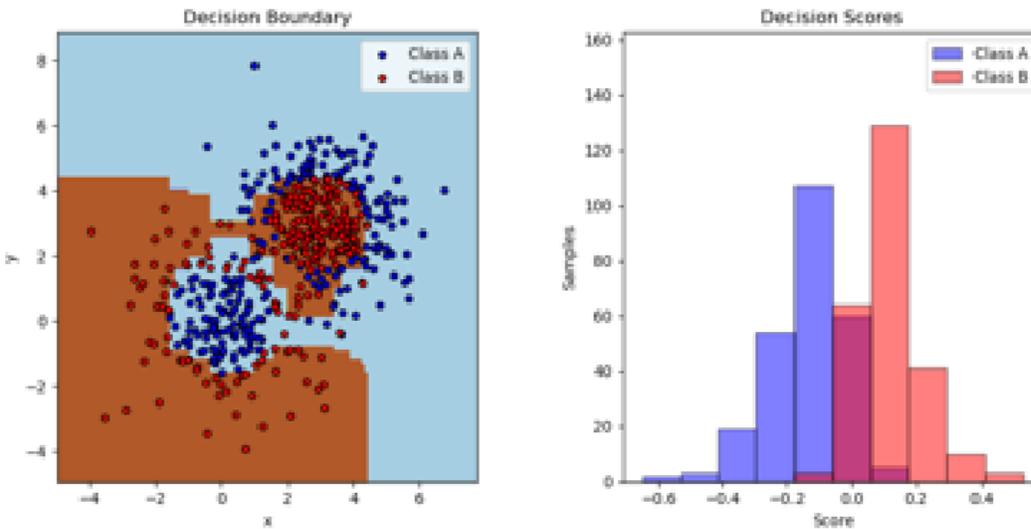
Key features:

- Personalized communication
- Risk mitigation
- Domain-specific research
- Discovery acceleration
- Interaction enrichment
- System condition monitoring
- Liabilities detection
- Disruption anticipation
- Chatbots & knowledge management
- Visual recognition & discovery news
- Speech to text & text to speech
- Natural language classifier
- Tone analyzer
- Personality insights
- Language translator & native language classifier

3 [Scikit-learn](#)

Scikit-learn is the free artificial intelligence tool that provides a variety of supervised and unsupervised learning algorithms through a consistent interface. It is considered a simple and efficient tool for data mining and data analysis. In various contexts, this free open-source AI software is reusable and accessible to everybody. You can consider this app if you want to bring machine learning into a production system.

This software serves as a convenient platform that can predict consumer behavior. Around the world, Scikit-learn is extensively used by commercial and research organizations. The business entities and research organizations find the module of this software easy to use as it allows performing the multitude of processes.



(Source: [Scikit-learn](#))

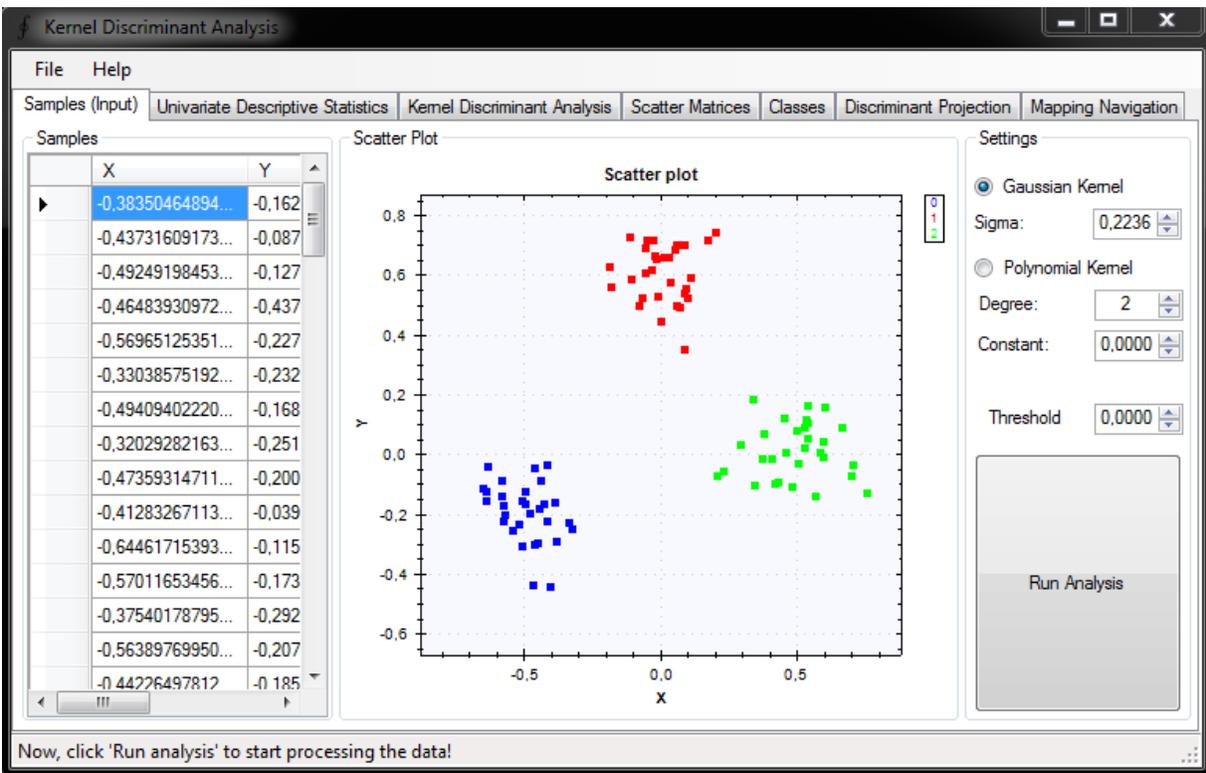
Key features:

- Adaptive & data ingestion
- Clustering & Classification
- Preprocessing
- Model Selection
- Dimensionality reduction
- Regression
- Cross-Validation
- Datasets & digital Assistants
- Dimensionality Reduction
- Manifold Learning
- Machine learning
- Process/Workflow Automation
- Virtual Personal Assistant (VPA)

4 [Accord.NET](#)

If you are looking for a free artificial intelligence tool, then Accord.NET is apt for your requirement. This app is a .NET machine learning framework that combines with audio and image processing libraries. This open-source AI development tool offers statistical analysis, image processing, machine learning, and mathematics and computer vision.

For creating production-grade computer vision, audition of computer, signal dispensation, statistics applications, and commercial use, you will find this software a complete framework. The structure of this software is divided into a library that includes scientific computing along with signal and image processing.



(Source: Accord.NET)

Key features:

- Data Ingestion
- Adaptive
- Support Vector Machines
- Deep learning
- Camshift object tracking
- Haar-feature image recognition
- Decision Trees along with automatic code generation
- Hypothesis Testing
- Statistical Analysis
- Hidden and conditional random fields
- Standard & multinomial logistic regression
- Second-order neural network learning algorithms
- Digital Assistants & Multi languages
- Process/Workflow Automation
- Speech recognition

5 Torch

This free, open-source artificial intelligence tool is a scientific computing framework with ample support for machine learning algorithms that places GPUs first. Torch believes in having complete flexibility and speed in building scientific algorithms. It makes the process enormously simple.

In machine learning, computer vision, parallel processing, signal processing, networking, image, video, audio, you will find that this free AI software app comes with a vast ecosystem of community-driven packages.

Consistently, this app is evolving and is used within Facebook, Twitter, NYU, Purdue, numerous research labs, and companies. It is swift and holds massive data along with new dynamic scoring.

Key features:

- Chatbot & speech recognition
- Adaptive & Digital Assistants
- N-dimensional array
- Routines for indexing, slicing transposing
- GPU support
- Fixed with ports to iOS and Android backends
- Neural network
- Energy-based models
- Predictive Analytics
- Machine learning

AI Software	Platform	Digital Assistants	Machine Learning	Adaptive	Data Ingestion	Chatbot	Predictive Analytics	Speech Recognition
Tensor Flow	Linux, macOS, Windows, Android, iPhone/iPad	✓	✓	✓	✓	✗	✓	✓
IBM Watson	Windows, Mac, Linux, Web-based	✓	✓	✓	✓	✗	✓	✓
Apache Mahout	Windows, iPhone/iPad, Web-based	✓	✗	✓	✗	✓	✗	✓
OpenNN	Windows, Android, Web-based, iPhone/iPad	✓	✗	✓	✗	✓	✓	✓
Scikit-Learn	Windows, Web-based	✓	✓	✓	✓	✗	✗	✗
accord.Net	Windows, Android, iPhone/iPad, web-based	✓	✗	✓	✓	✗	✗	✓
Torch	Windows, android, iPhone/iPad, web-based	✓	✓	✓	✗	✓	✓	✓

Figure 2 Comparison chart of Artificial Intelligence Software

NETWAK datasets and prototype training on robust, scalable Artificial Intelligence Natural Language Planning innovation technologies provide the next generation invention in data identification, data authentication, data recording and analysis into information, information synthesis, retrieval, rechecks, control and reporting (Chhabra & Suri, 2019).

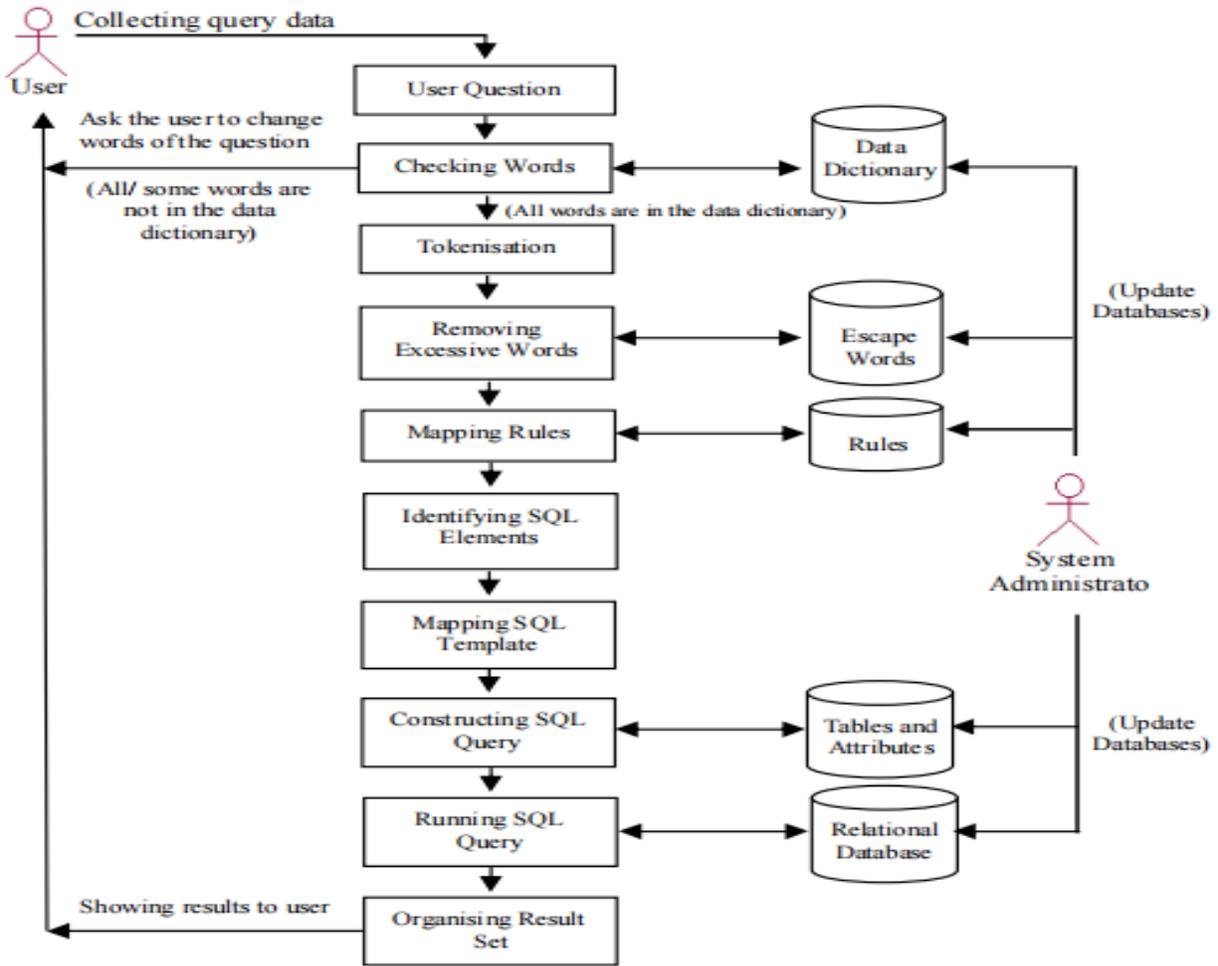


Figure 3 NETWAK DATA COLLECTION, PROCESSING AND CROSS REFERENCE DATABASE CREATION

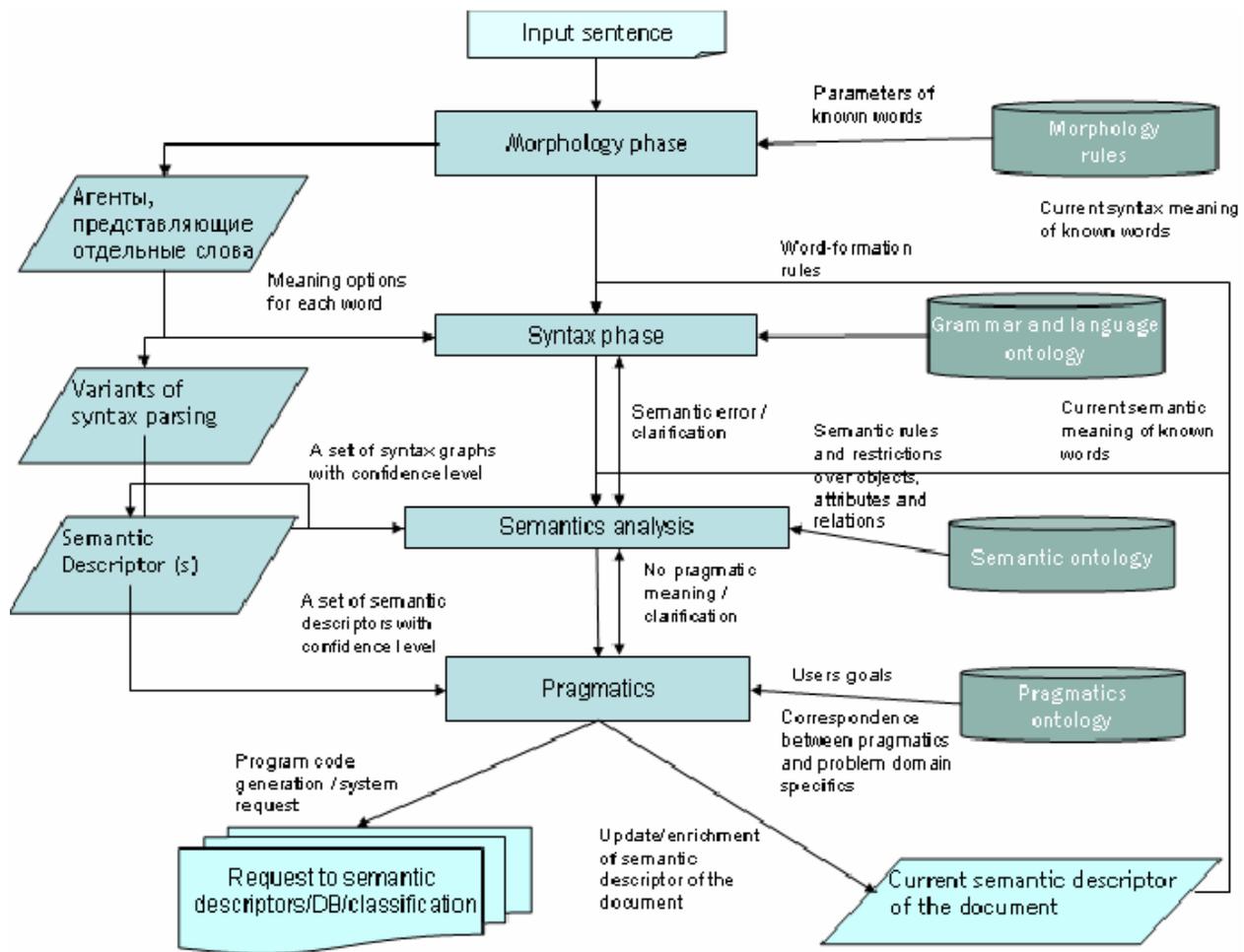


Figure 4 NATURAL LANGUAGE PROCESSING ALGORITHM

Care and due diligence should be taken in choosing the right mix of open source, cheap or free API software, cloud-based technologies, and programmed coded own set of rules for machine learning algorithms and software applications.

For example, a set of rules are set up in programming to include each phase in the above natural language processing algorithm diagram in developing recognized tokens and well-structured pieces of sentences to become a set of machine learning rules that evolve on their own experience with data collected over time (Zhou, 2022).

Cross referencing to determine similarity in future user inputs determine the authenticity of user requests in reference to the self-evolving and continuously growing database (Li, 2011).

NETWAK ACTIVITIES

NETWAK networks include communication and collaboration with country government education ministries, resource websites concerned with school-based information, free and open access academic and educational materials and courses like MOOCS – Massive Open Online Courses, and other OER Open Educational Resources institutions.

Some of corresponding institutions include <https://www.oeconsortium.org/>, <https://oeru.org/>, <https://sparcopen.org/open-education/>, and other similar organizations. The main objective of NETWAK should not be lost in catering for the disadvantaged student who cannot even cheap access to online education.

In that respect it is the endeavor of NETWAK to scale and integrate digital radio and television education classes organized with the collaborating institutions offering free global education, as to reach the most disadvantaged at the village grassroots level mostly in developing global country governments.

NETWAK systems and processes are aimed at delivering free and open access of education at all levels to undeserving and unable students globally. To achieve this objective, network activities are geared towards the implementation of affordable technological solutions to enable communication with such students that have little or no access to education due to financial constraints.

The NETWAK system identifies, sources and authenticates the genuine applications from needy students and matches them with relevant certifiable open access courses, scholarships and funding from benevolent and philanthropic organizations worldwide.

The NETWAK identifies needy students by synthesizing information recognition metrics into deeply embedded ED-TECH systems. Background checks on needy student requisitions for education help is an intrinsic part of in-country NETWAK administrative systems that report to head office.

CONCLUSION

Implementing Artificial Intelligence NLP Natural Language Processing network topologies and model designs in NETWAK build, takes advantage of the next generation deep learning neural networking secret that will spur industries to scale to outcompete their competitors because of massive cost reductions when scaling prototypes globally due to the utilization of free and open-source resources.

Building up a corpus relational referencing database for voice, text and pictorial search inputs for NETWAK websites and software applications not only creates an efficient authentication and security audit system but also streamlines NETWAK systems to a unique robust database engine that scales above other similar or related educational networking organizations, platforms and services globally.

INDEX

Related Collaboration network websites;

<https://univ.cc/search.php?dom=world&key=&start=1>

<https://www.edx.org/>

<https://www.opencolleges.edu.au/informed/features/free-online-courses-50-sites-to-get-educated-for-free/>

<https://www.guru99.com/free-online-education-sites.html>

<https://www.topuniversities.com/student-info/distance-learning/9-top-universities-offering-free-online-courses>

www.scholarship-positions.com

<https://www.opportunitiesforafricans.com/category/scholarships/>

<https://www.topuniversities.com/student-info/scholarship-advice/international-scholarships-african-students>

<https://www.advance-africa.com/Scholarships-for-Study-in-Africa.html>

<https://www.scholars4dev.com/category/target-group/asians-scholarships/>

<https://www.topuniversities.com/student-info/scholarship-advice/international-scholarships-study-asia>

<https://www.afterschoolafrica.com/scholarship/by-country/scholarship-in-asia/>

<https://www.scholarshipportal.com/scholarships/asia>

<https://www.scholars4dev.com/category/target-group/latin-americans-scholarships/>

<https://www.scholars4dev.com/category/target-group/latin-americans-scholarships/>

<https://www.scholars4dev.com/category/target-group/latin-americans-scholarships/>

References

- Chhabra, I., & Suri, G. (2019). Knowledge Discovery for Scalable Data Mining. *ICST Transactions on Scalable Information Systems*, 6(21), 158527. <https://doi.org/10.4108/eai.19-3-2019.158527>
- Goldberg, Y. (2016). A Primer on Neural Network Models for Natural Language Processing. *Journal of Artificial Intelligence Research*, 57(1510.00726), 345–420. <https://doi.org/10.1613/jair.4992>
- Li, H. (2011). Learning to Rank for Information Retrieval and Natural Language Processing. *Synthesis Lectures on Human Language Technologies*, 4(1), 1–113. <https://doi.org/10.2200/s00348ed1v01y201104hlt012>
- Vasilenko, K., Zolkin, A., Abramov, N., & Kurganov, D. (2020). ENTERPRISE DATA PROCESSING NETWORKS: AUTHENTICATION AND IDENTIFICATION PROCEDURES. *Infokommunikacionnye Tehnologii*, 18(No. 1 (2020)), 61–67. <https://doi.org/10.18469/ikt.2020.18.1.10>
- Zhou, Y. (2022). Natural Language Processing with Improved Deep Learning Neural Networks. *Scientific Programming*, 2022, 1–8. <https://doi.org/10.1155/2022/6028693>