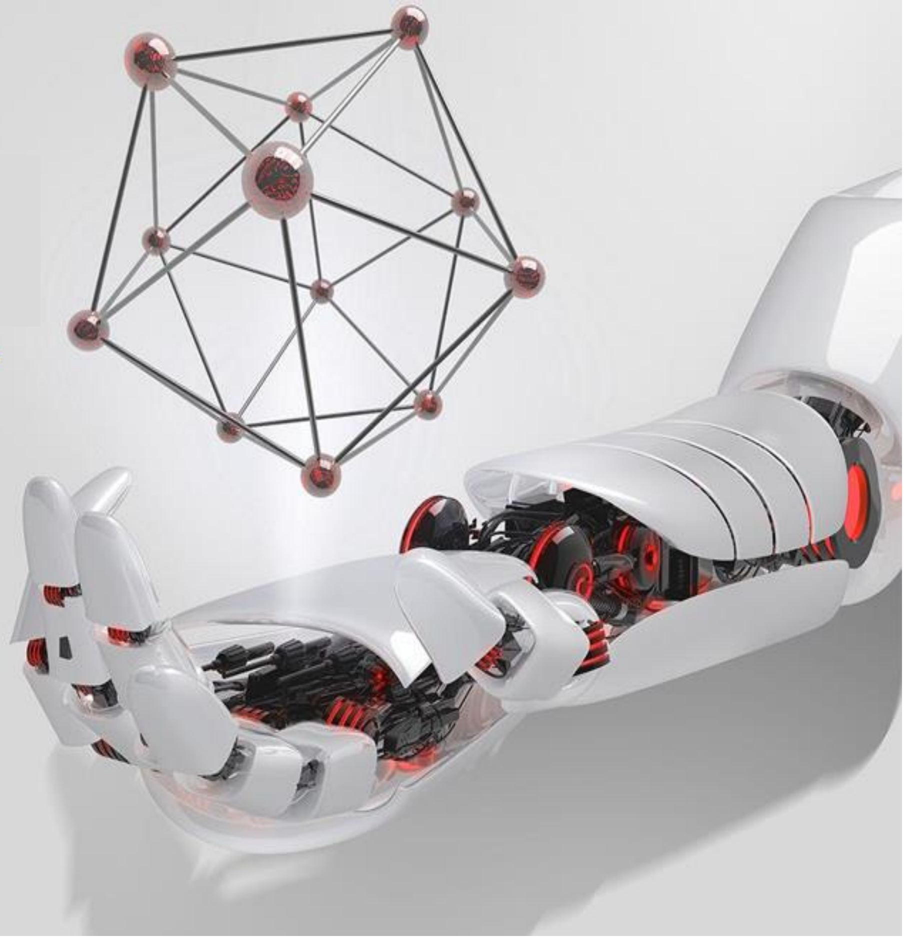
IDEAS WORTH SPREADING

Language Models, Intents and Contexts:

Domain-Aware Text Understanding Simplified



Dr. Sourav Dutta

Chief NLP Research Scientist

Huawei Ireland Research

NLP Research

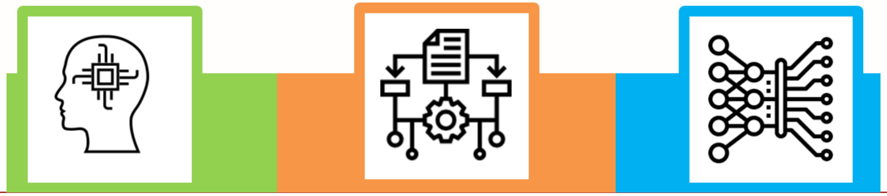
** Figures used in this presentation have been obtained via Internet search results with / without modifications.

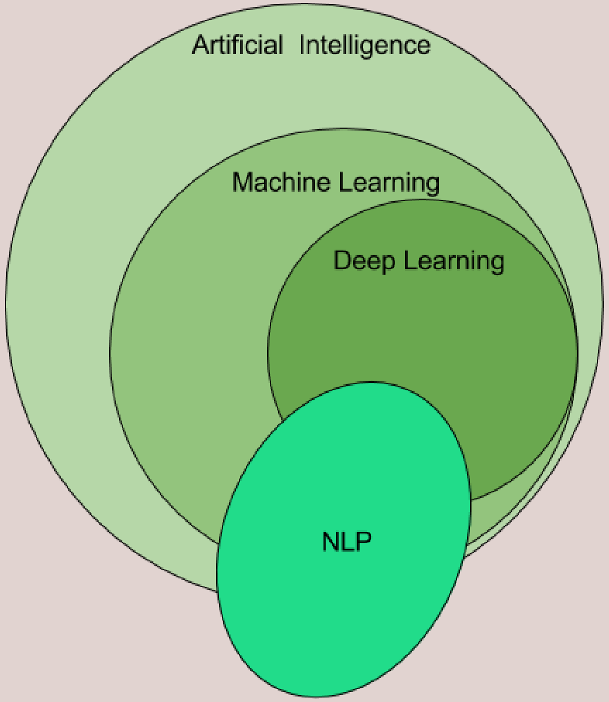
The Splurge of Data

Staggering amounts of structured and unstructured textual data

- **700K** search queries
- **100K** Tweets
- **20K** posts on Tumblr
- **12K** ads posted
- **2K** blogs created
- Questions & Answers
- Streams of News Articles

NLP@Huawei





NLP team focuses on researching and developing the next generation of multilingual NLP technologies for improving end user experience.

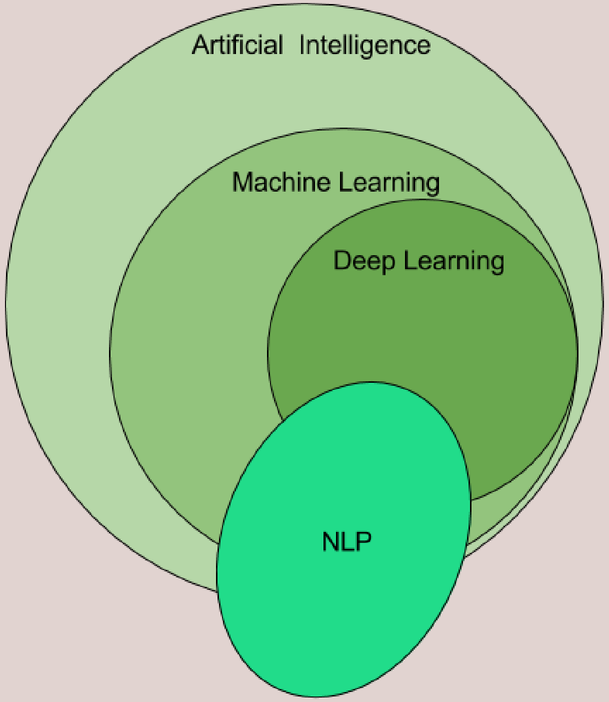
- **Vision**

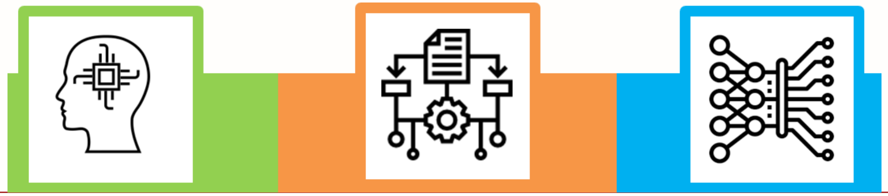
  Building world-class NLP capability by focusing on fundamental research and building novel cutting-edge core NLP framework that can be utilized across various HMS services.

- **Mission**

  End-to-end multilingual NLP algorithmic stack.

NLP@Huawei





NLP team focuses on researching and developing the next generation of multilingual NLP technologies for improving end user experience.

- **Vision**

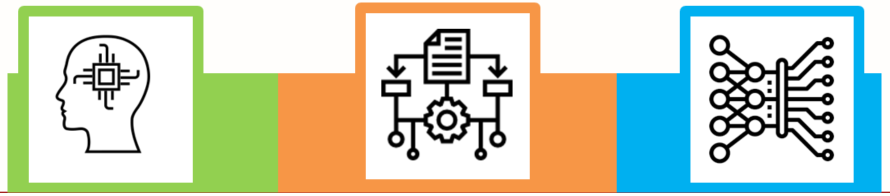
  Building world-class NLP capability by focusing on fundamental research and building novel cutting-edge core NLP framework that can be utilized across various HMS services.

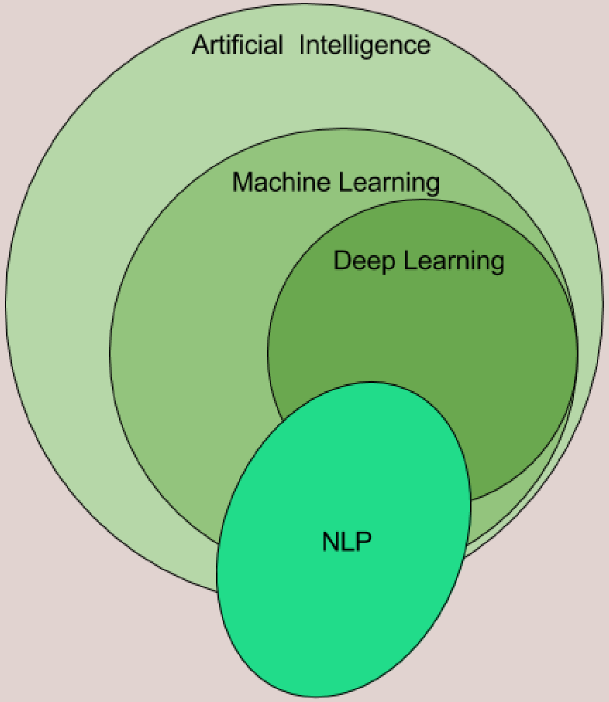
- **Mission**

  End-to-end multilingual NLP algorithmic stack.

- **Research Topics**
  - *Multi-Lingual Text Embedding*
    - New text representation framework to capture enhanced semantic understanding across various languages for better search, classification and zero-shot transfer capabilities.
  - *Question Answering*
    - Intelligent chatbot solution for understanding and catering to users' diverse informational need across diverse domains for different applications.

NLP@Huawei





NLP team focuses on researching and developing the next generation of multilingual NLP technologies for improving end user experience.

- **Vision**

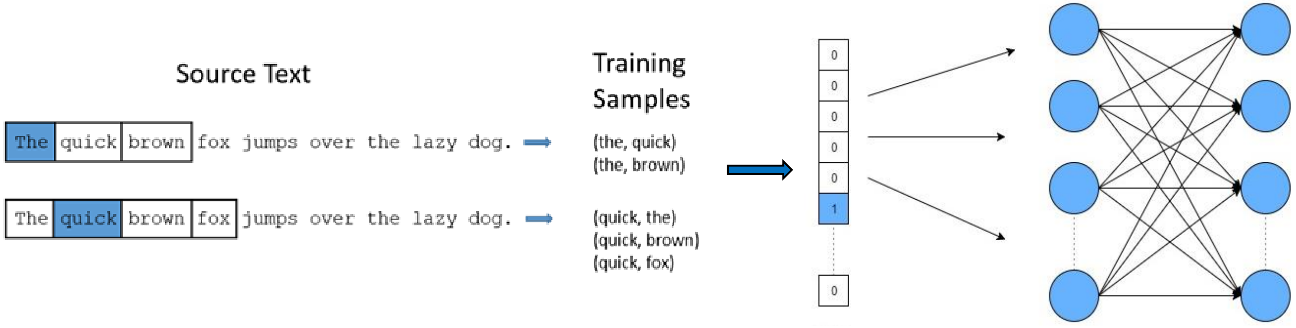
  Building world-class NLP capability by focusing on fundamental research and building novel cutting-edge core NLP framework that can be utilized across various HMS services.

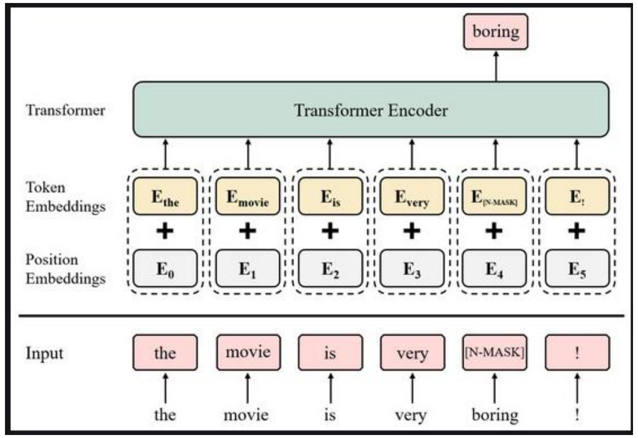
- **Mission**

  End-to-end multilingual NLP algorithmic stack.

- **Research Topics**
  - *Multi-Lingual Text Embedding*
    - New text representation framework to capture enhanced semantic understanding across various languages for better search, classification and zero-shot transfer capabilities.
  - *Question Answering*
    - Intelligent chatbot solution for understanding and catering to users' diverse informational need across diverse domains for different applications.
  - *Knowledge Graph based Learning*
    - Automated creation of Knowledge Graphs and their integration with context for better understanding of semantics and relationships within text.
  - *Neural Network Pruning*
    - Compression of language models for smaller but accurate textual understanding.

Context as Vector Embedding

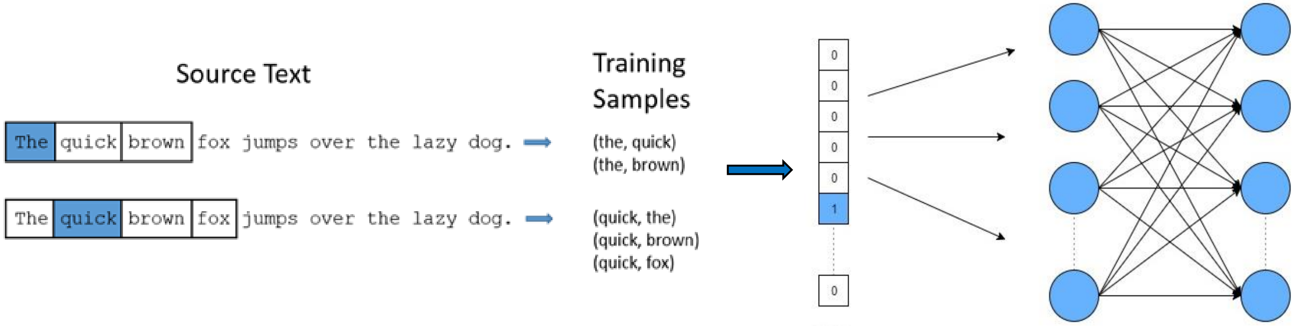


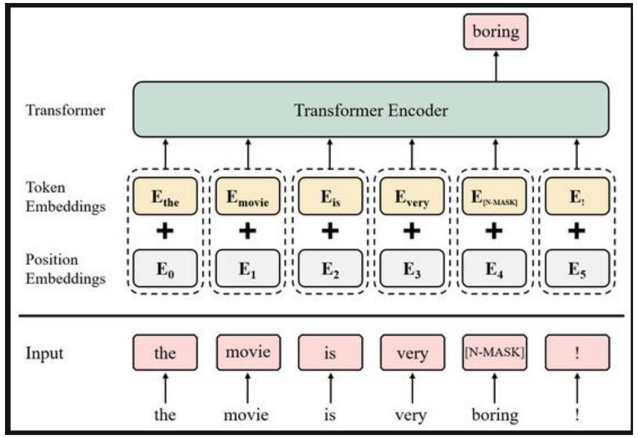


- **Captures word, semantic and concept from the context of texts.**
  - Learns about the "distributional hypothesis" of words.
  - Provides a vector representation *(embedding)* of the word.
- **Mathematical operations on words.**
  - Berlin – Germany + France ≈ Paris
- BERT, XLM-R, InfoXLM, etc.

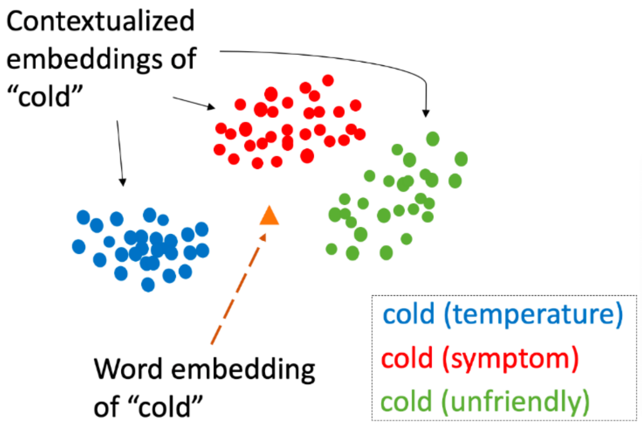
T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013). "Distributed representations of words and phrases and their compositionality". Advances in Neural Information Processing Systems.
J. Pennington, R. Socher and C. D. Manning (2014). "GloVe: Global Vectors for Word Representation". EMNLP.

Context as Vector Embedding





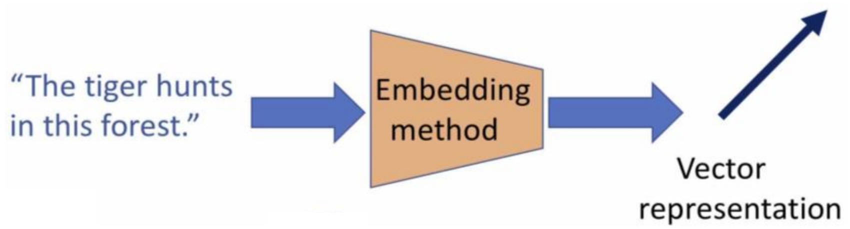
- **Captures word, semantic and concept from the context of texts.**
  - Learns about the "distributional hypothesis" of words.
  - Provides a vector representation *(embedding)* of the word.
- **Mathematical operations on words.**
  - Berlin – Germany + France ≈ Paris
- BERT, XLM-R, InfoXLM, etc.



T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013). "Distributed representations of words and phrases and their compositionality". Advances in Neural Information Processing Systems.
J. Pennington, R. Socher and C. D. Manning (2014). "GloVe: Global Vectors for Word Representation". EMNLP.

Embedding Sentences & Multi-Linguality

- Cross-Lingual Sentence Embedding are an important component for *Natural Language Understanding* tasks.
  - Search and Information Retrieval
  - Semantic Understanding of user intent
- **Sentence Embedding**: *dense vector representations encoding contextual and semantic information.*
  - Sentences of *similar meanings* will have *close representations*

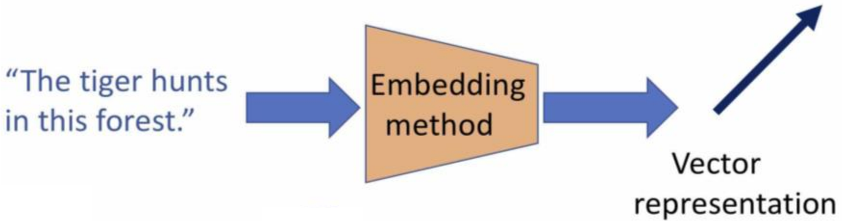


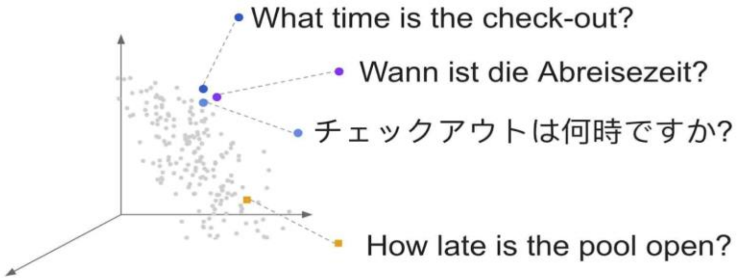
N. Reimers and I. Gurevych (2019). "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. ACL.
D. Cer et al. (2018). Universal Sentence Encoder.
A. Conneau et al. (2019). "Unsupervised Cross-Lingual Representation Learning at Scale".
M. Artetxe and H. Schwenk (2018). "Massively Multilingual Sentence Embeddings for Zero-Shot Cross-Lingual Transfer and Beyond."

Embedding Sentences & Multi-Linguality

- Cross-Lingual Sentence Embedding are an important component for *Natural Language Understanding* tasks.
    - Search and Information Retrieval
    - Semantic Understanding of user intent
- **Sentence Embedding**: *dense vector representations encoding contextual and semantic information.*
    - Sentences of *similar meanings* will have *close representations*
- **Multi-Lingual Sentence Embedding:** provide cross-lingual representation of text across languages.
    - Efficient way of mapping multilingual texts into a unified vector space.

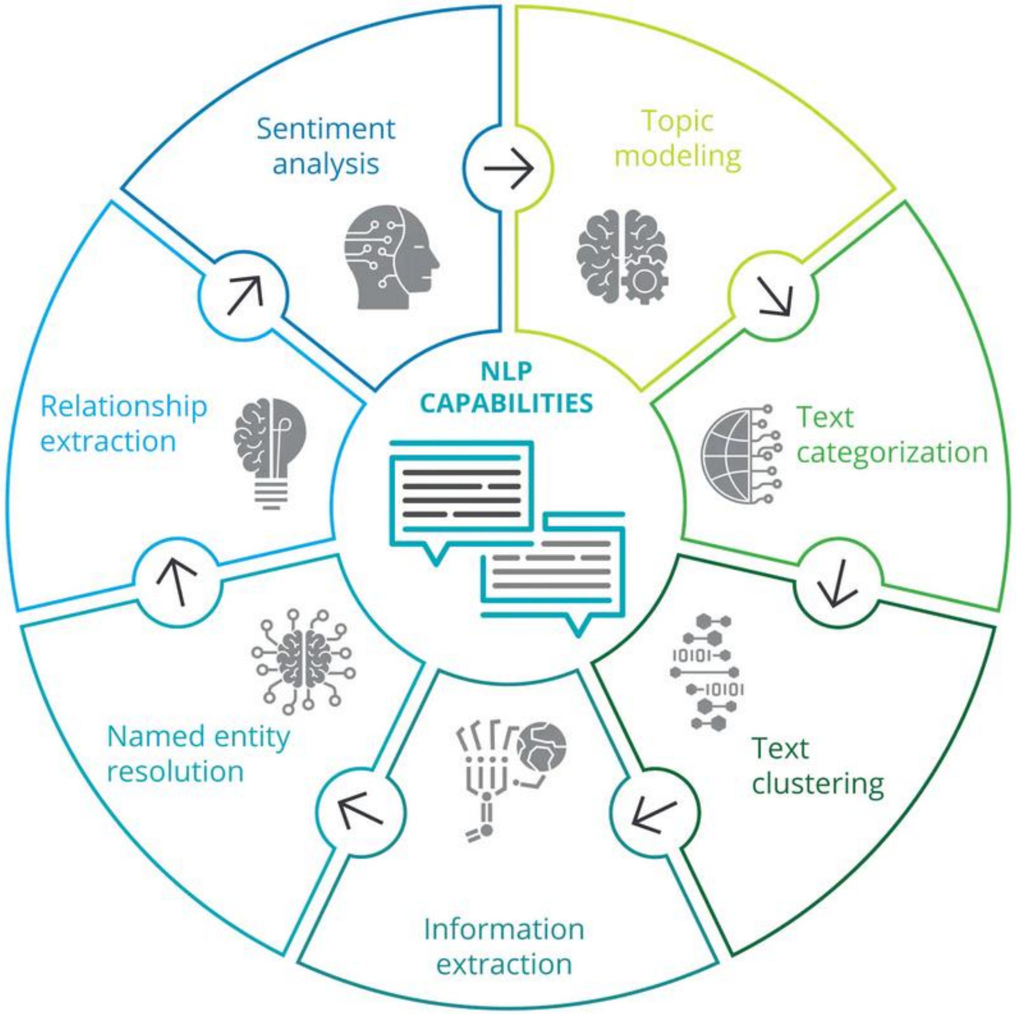
- Sentence Transformers (S-BERT), Universal Sentence Encoder (USE), LASER, LaBSE etc.





N. Reimers and I. Gurevych (2019). "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. ACL.
D. Cer et al. (2018). Universal Sentence Encoder.
A. Conneau et al. (2019). "Unsupervised Cross-Lingual Representation Learning at Scale".
M. Artetxe and H. Schwenk (2018). "Massively Multilingual Sentence Embeddings for Zero-Shot Cross-Lingual Transfer and Beyond."

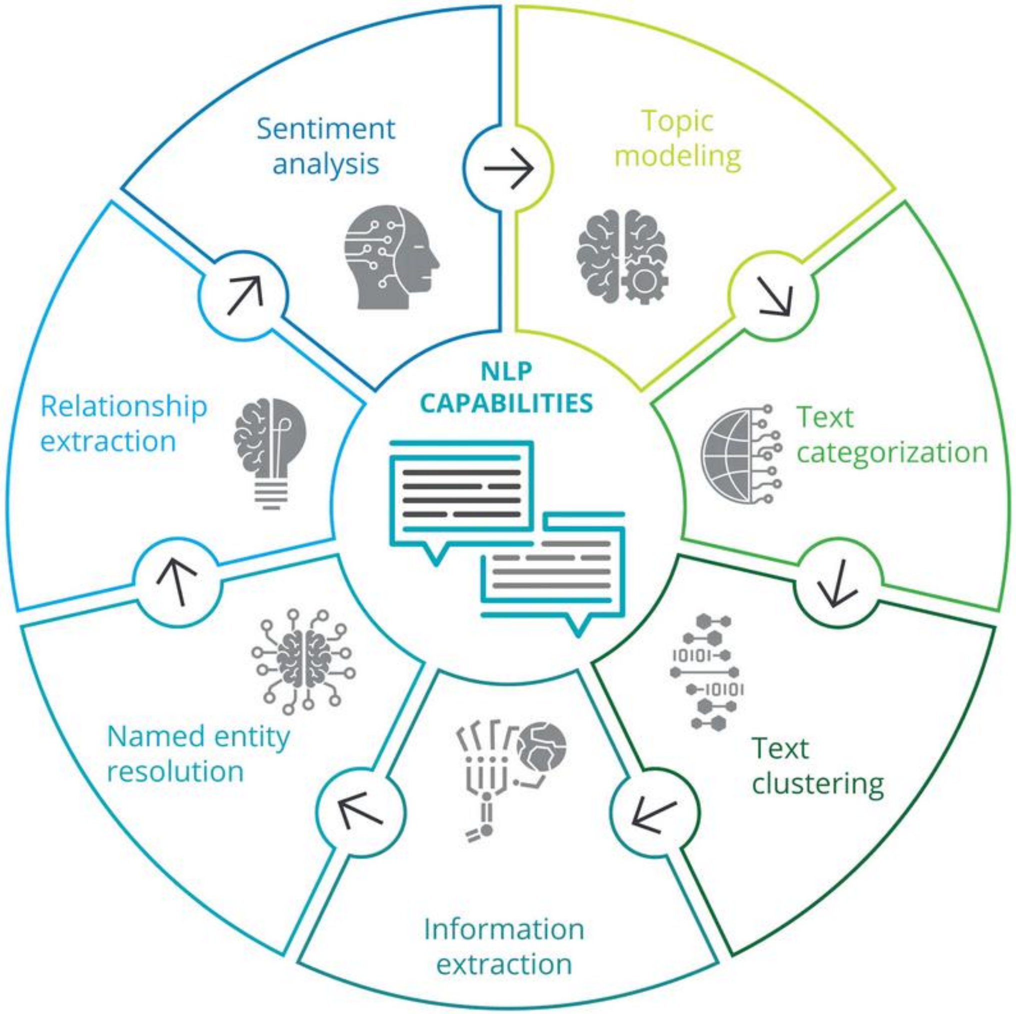
Application Scenarios

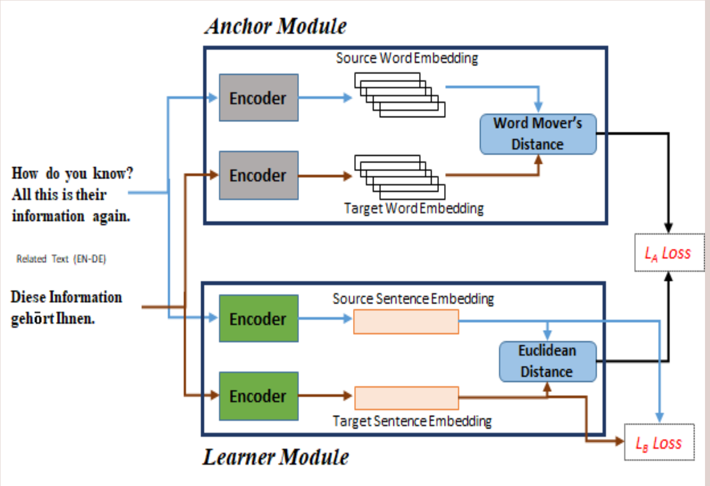


Application Scenarios

- Creditworthiness assessment
- Advertising and Marketing
- Hiring and recruitment
- Sentiment Analysis
- Chatbots
- Fake news identification

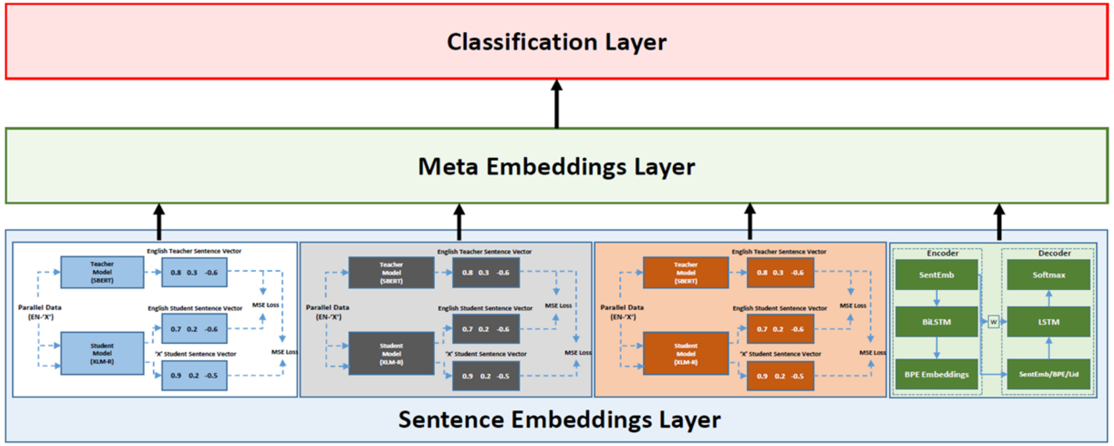
Machine Translation
Automatic Summarization/ Paraphracing
Natural Language Generation
Reasoning over Knowledge Based
Quation Answering System
Dialog System
Image Captioning & other Multimodel Tasks





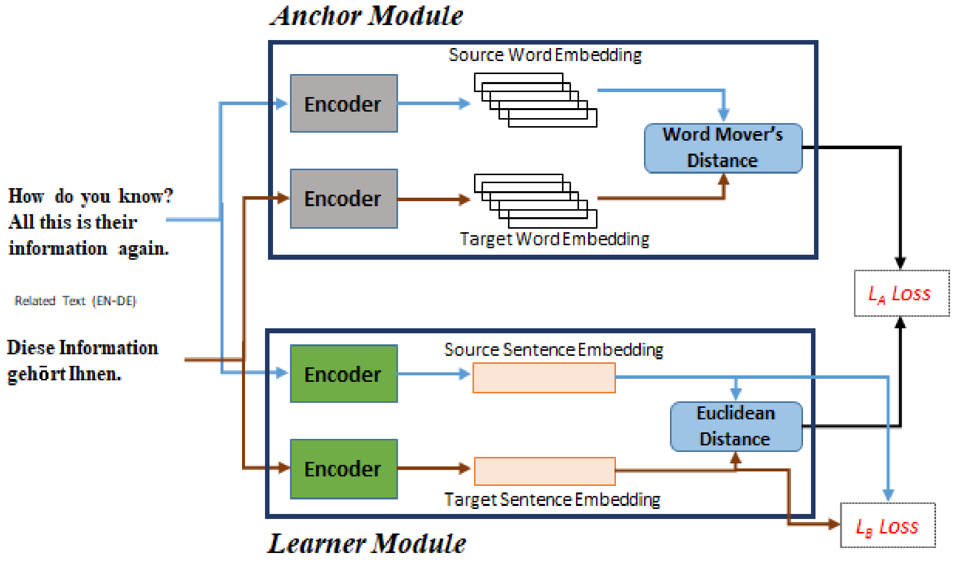
SEMANTIC UNDERSTANDING OF TEXT

The Curious Case of User Intent Understanding



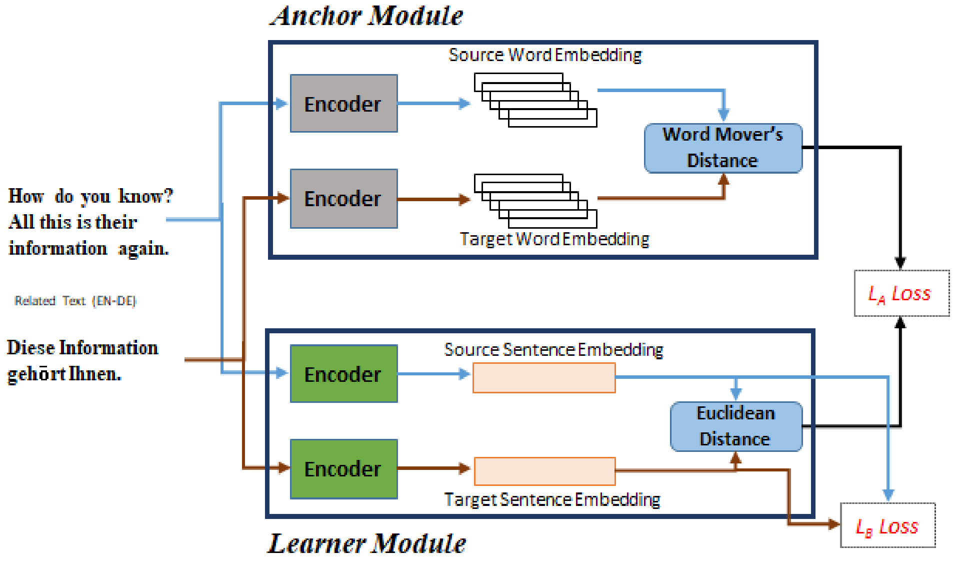
Dual Encoder based Sentence Embedding

- A novel *dual-encoder* based *anchor-learner* architecture.
- A joint loss function coupling *Word Mover's Distance* and *Cosine similarity*.



Dual Encoder based Sentence Embedding

- A novel *dual-encoder* based *anchor-learner* architecture.
- A joint loss function coupling *Word Mover's Distance* and *Cosine similarity*.
- *Learner* generates sentence embedding in *multi-task* setup
  - *Unsupervised* Loss function L_A captures *semantic* relationship between sentence pairs
    - Close proximity for similar sentences in different languages
  - Loss function L_B helps to map correct *translation pairs*
    - Transfer of cross-lingual linguistic properties



$$\mathcal{L}_{\mathcal{A}} = \frac{1}{N}\sum_{i=1}^{N} \exp^{|\exp^{-d_{euc}(s'_i, t'_i)} - exp^{-d_{wmd}(s_i, t_i)}|}$$

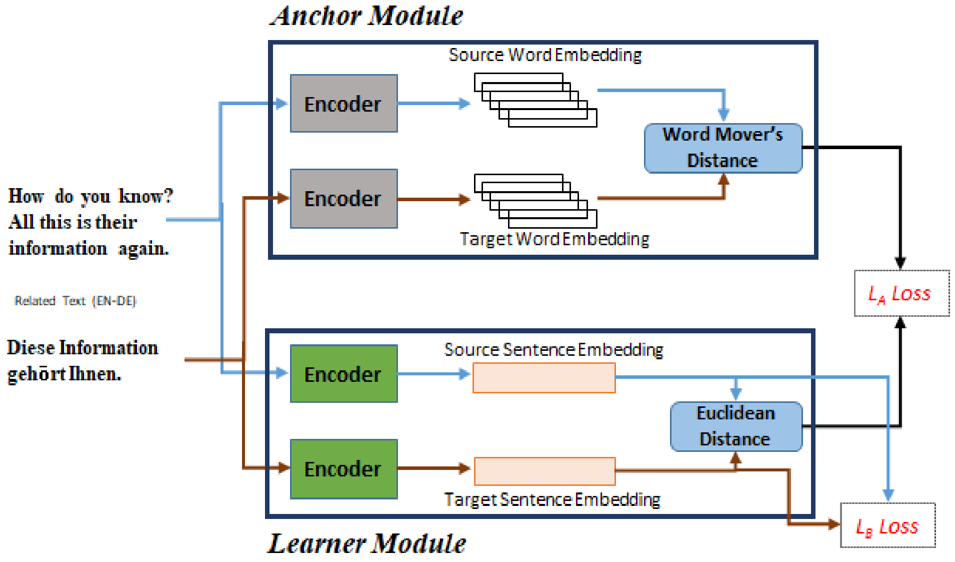
$$\mathcal{L}_{\mathcal{B}} = \frac{1}{N}\sum_{i=1}^{N} cossim(s'_i, t'_i)$$

Dual Encoder based Sentence Embedding

- A novel *dual-encoder* based *anchor-learner* architecture.
- A joint loss function coupling *Word Mover's Distance* and *Cosine similarity*.

- *Learner* generates sentence embedding in *multi-task* setup
  - *Unsupervised* Loss function L_A captures *semantic* relationship between sentence pairs
    - Close proximity for similar sentences in different languages
  - Loss function L_B helps to map correct *translation pairs*
    - Transfer of cross-lingual linguistic properties

- No requirement of parallel corpus.
- Robustness in performance for *Low-Resource Languages*

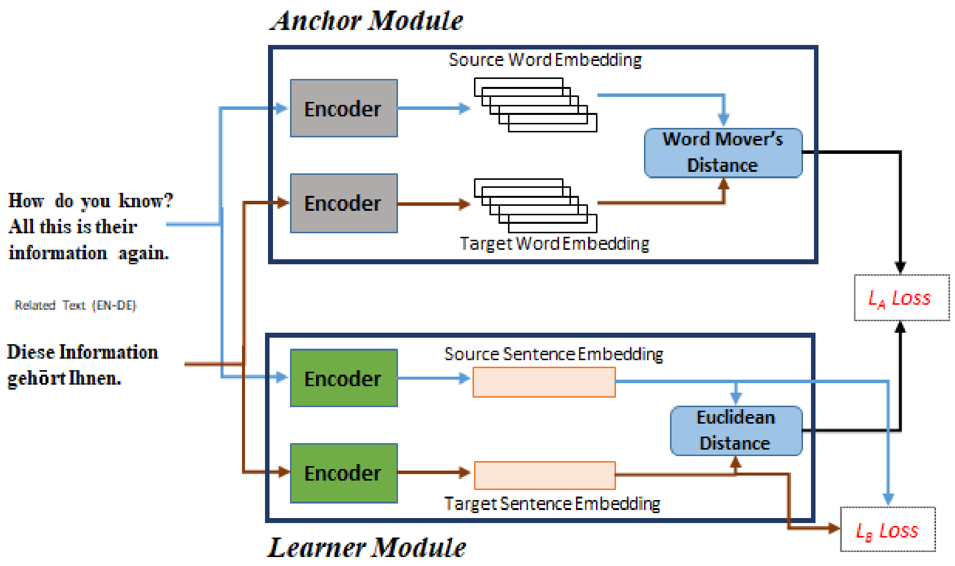


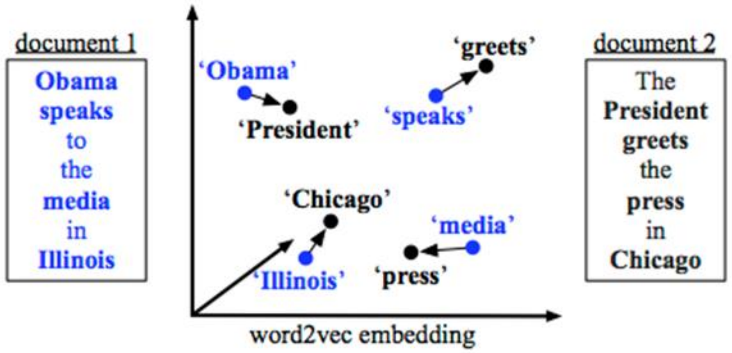
$$\mathcal{L}_{\mathcal{A}} = \frac{1}{N}\sum_{i=1}^{N} \exp^{\left|\exp^{-d_{euc}(s'_i,t'_i)} - exp^{-d_{wmd}(s_i,t_i)}\right|}$$

$$\mathcal{L}_{\mathcal{B}} = \frac{1}{N}\sum_{i=1}^{N} cossim(s'_i, t'_i)$$

Dual Encoder based Sentence Embedding

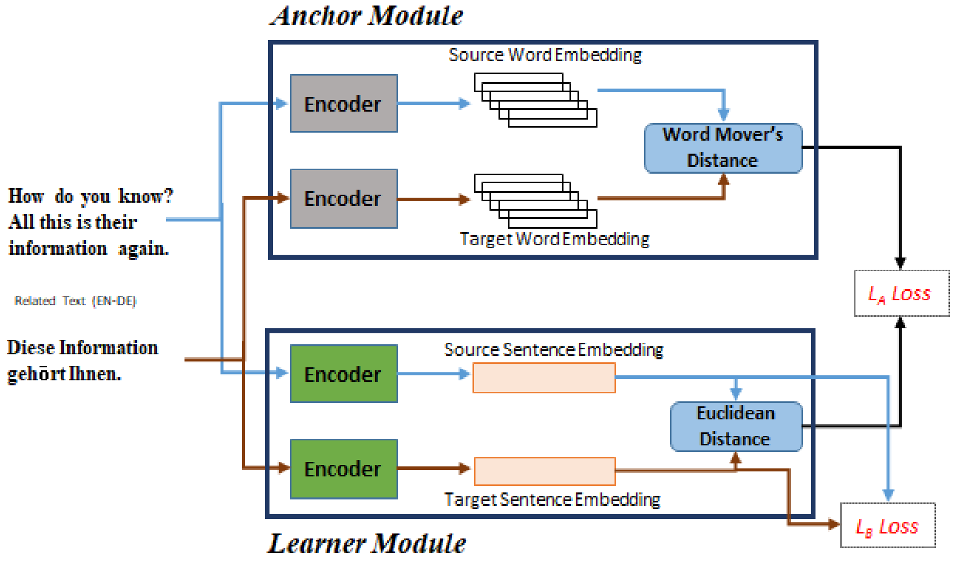
- Word Mover's Distance provides an optimal transportation formulation (adapted from Earth Mover's Distance).
  - The effort to transform the word representations between the input documents.





M. Kusner, Y. Sun, N. Kolkin, and K. Weinberger (2015). "From Word Embeddings To Document Distances. ICML.

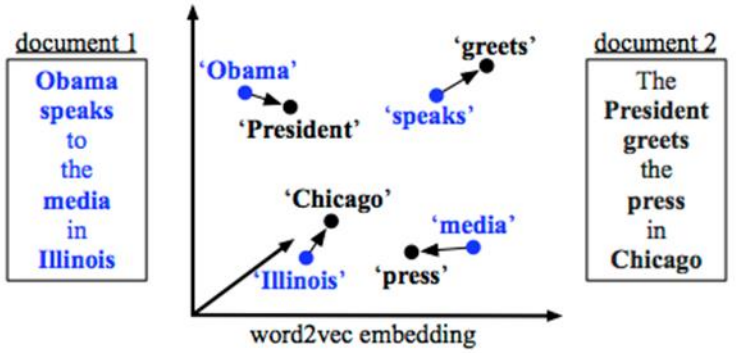
Dual Encoder based Sentence Embedding



- Word Mover's Distance provides an optimal transportation formulation (adapted from Earth Mover's Distance).
  - The effort to transform the word representations between the input documents.

- Inclusion of *Word Mover's Distance* is *advantageous* for *unsupervised* learning – *a proxy for semantic similarity*.
  - Closer representations for similar sentences.
  - Dissimilar sentences have embedding that are apart in embedding space.

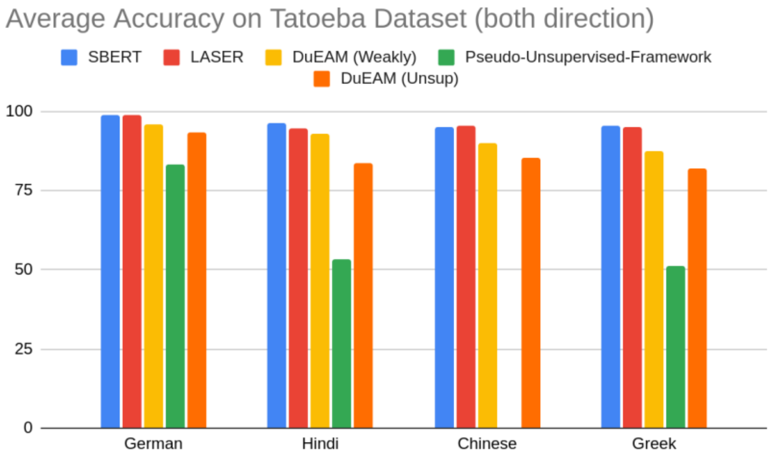
$$\mathcal{L}_A = \frac{1}{N}\sum_{i=1}^{N} \exp^{|\exp^{-d_{euc}(s'_i,t'_i)} - exp^{-d_{wmd}(s_i,t_i)}|}$$



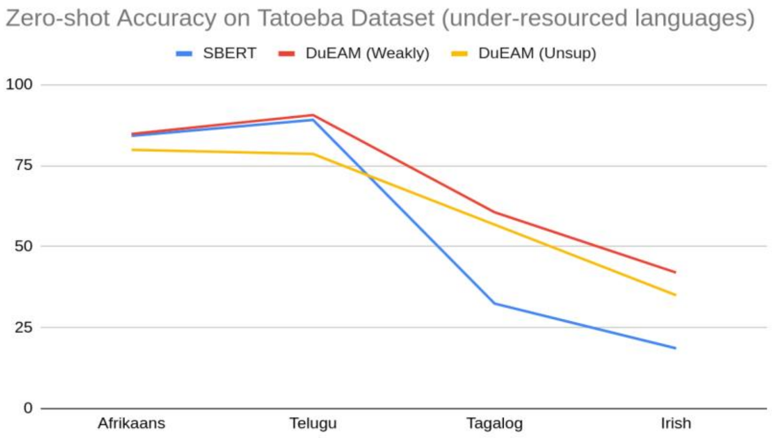
M. Kusner, Y. Sun, N. Kolkin, and K. Weinberger (2015). "From Word Embeddings To Document Distances. ICML.

Experimental Validation

Unsupervised Methods						
mBERT mean	54.4	56.7	33.9	16.0	21.5	33.0
XLM-R mean	50.7	51.8	21.3	9.2	10.9	16.6
Proposed methods						
DuEAM$_{wklysupv}$	**81.9**	**83.1**	**69.4**	**68.6**	**64.6**	**69.6**
DuEAM$_{unsupv}$	*80.2*	*81.5*	*64.6*	*63.7*	*58.2*	*62.1*

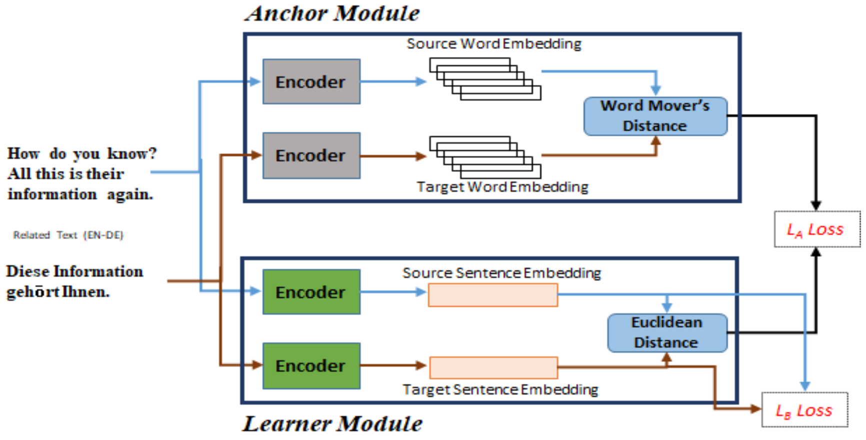


- *State-of-the-art Unsupervised* Sentence Embedding performance
- Proposed models produced *better monolingual* results than *LaBSE* and *LASER.*
- *Weakly-supervised* model produced *comparative results* to *supervised* models for high-resourced languages.
- Results in *zero-shot* settings is eye catching.



Text Embedding – In a nutshell

- Unsupervised and Weakly-Supervised Multilingual Sentence Embedding Model based on anchor-learner framework.
- Word Mover's Distance based unsupervised loss function captures semantic relatedness efficiently between sentence pairs.
- State-of-the-art results on downstream Natural Language Processing tasks.
- Easy to extend for low-resourced languages.



Cross-lingual Sentence Embedding using Multi-Task Learning

Koustava Goswami[1*], Sourav Dutta[2], Haytham Assem[2], Theodorus Fransen[1] and John P. McCrae[1]

[1] Data Science Institute, National University of Ireland Galway, Ireland
[2] Huawei Research Centre, Dublin, Ireland
{koustava.goswami,theodorus.fransen, john.mccrae}@insight-centre.org
{sourav.dutta2, haytham.assem}@huawei.com

Text Classification – Automated Chatbot

- FAQs provide a collection of Question-Answer (QA) pairs
  - “One-stop” information source for users on products and services
  - Curated and well-defined by experts
- Natural Language interface for user search.
  - Automated customer service chatbots
- User question matching with short text span for relevant answers.
  - Semantically similar retrieval – “*Account login failed*” versus “*Unable to access my account*”

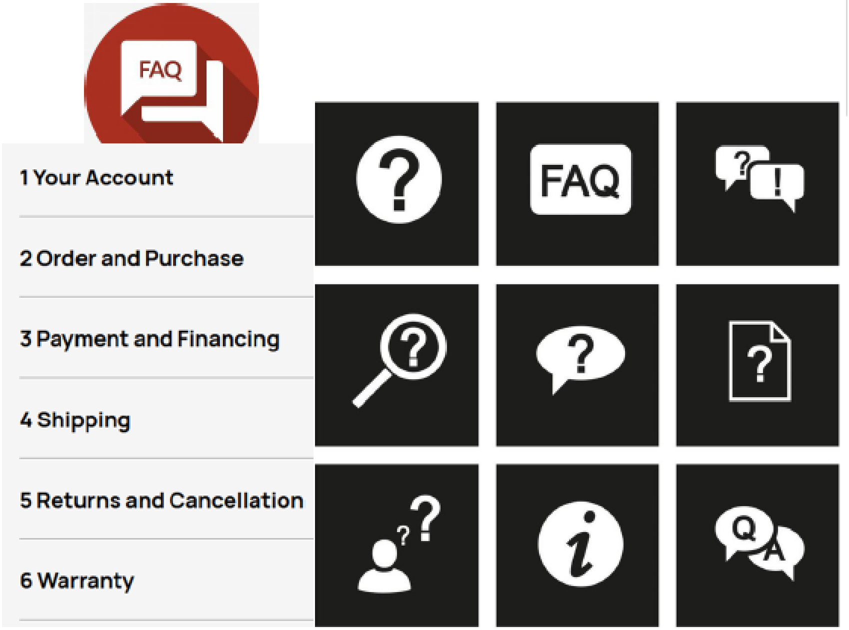


Figure 1. FAQ for user query answering.

Text Classification – Automated Chatbot

- The current trend is to fine-tune language models (BERT, XLM, etc.) or its variants for downstream multilingual sentence or short text classification tasks.

- But fine-tuning such large pre-trained language models is highly resource-intensive.
  - Limited training instances
  - Zero-shot transfer

- We present a practical and efficient framework based on fusing various pre-trained sentence encoders for ***text classification.***

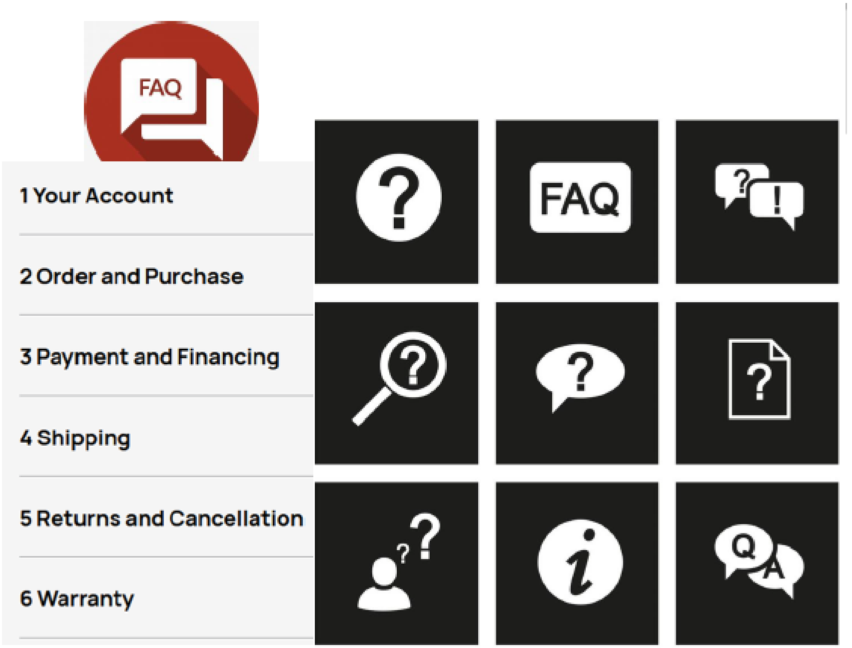
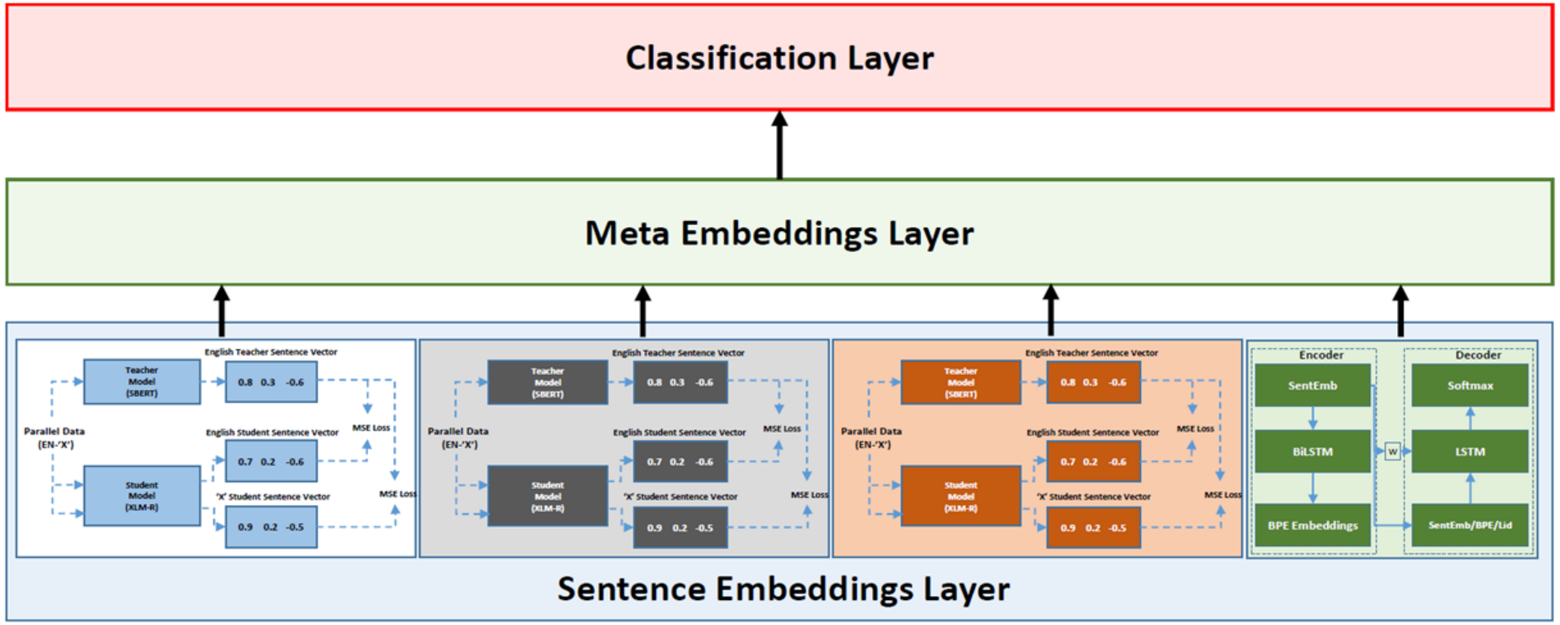


Figure 1. FAQ for user query answering.

Text Classification

- We borrow from the concept of "meta-embedding".
  - Multiple language models are used to understand different linguistic facets learnt by the different architectures.
  - *Concatenation of embedding*



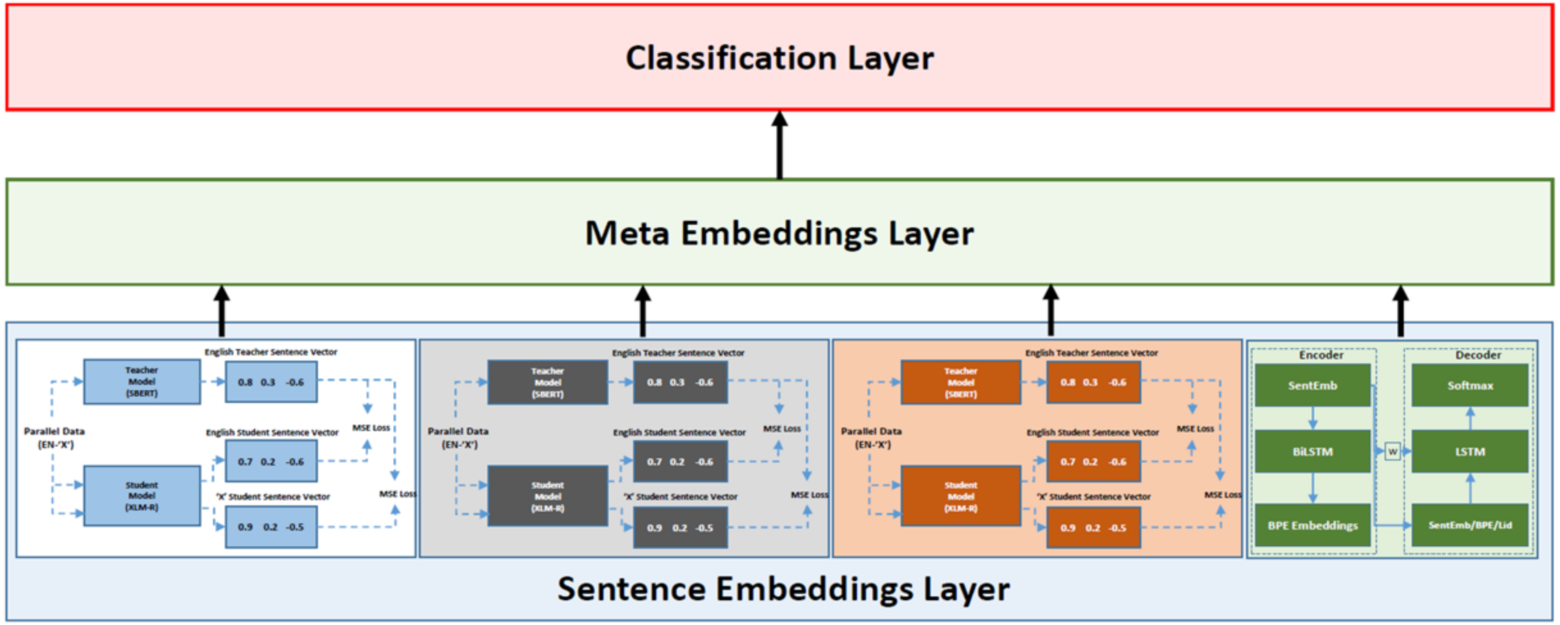
Multilingual Pre-trained Sentence Encoders

Text Classification

- We borrow from the concept of "meta-embedding".
    - Multiple language models are used to understand different linguistic facets learnt by the different architectures.
    - *Concatenation of embedding*

- We stack a shallow Multi-Layer Perceptron (***MLP***) with two/three hidden layer.

- This module computes the ***cross-entropy loss*** for the downstream classification tasks.



Multilingual Pre-trained Sentence Encoders

The cross-entropy loss is minimized so as to increase the **inter-class separability,** and **intra-class compactness**

Evaluation on Text Classification

INTENT UNDERSTANDING RESULTS ON *NLU* DATA: ZERO-SHOT

Baseline	*en*	*de*	*es*	*ar*	*hi*	*zh*	Avg.	Training Time (hr)	Inference Time (msec)
XLM-R-large	91.4	87.4	85.9	81.3	80.5	83.7	85.03	5.834	1.76
XLM-R-base	90.2	80.8	81.3	67.7	70.4	78.7	78.19	0.762	0.59
mBERT-base	88.9	56.2	61.7	37.4	30.2	57.3	55.28	0.336	0.62
DistilBERT	87.5	50.1	56.2	28.4	18.9	45.6	47.78	0.044	0.34
LASER	88.5	83.6	84.4	78.6	72.3	63.8	78.53	0.025	0.99
SBERT	87.4	83.4	83.6	78.8	81.2	82.7	82.85	0.025	1.07
Proposed	90.7	86.2	87.3	80.7	84.9	84.4	85.70	0.025	2.90

INTENT UNDERSTANDING RESULTS ON *MTOP* DATA: ZERO-SHOT

Baseline	*en*	*de*	*es*	*fr*	*hi*	*th*	Avg.	Training Time (hr)	Inference Time (msec)
XLM-R-large	96.3	92.3	92.2	91.2	89.1	89.7	91.8	6.080	1.80
XLM-R-base	96.6	88.4	88.0	86.1	79.6	85.9	87.43	0.835	0.59
mBERT-base	95.9	58.3	67.2	65.3	37.9	33.6	59.70	0.682	0.63
DistilBERT	95.3	53.0	61.9	59.8	31.6	22.1	53.95	0.368	0.33
LASER	94.9	91.9	89.6	90.1	80.9	88.3	89.28	0.119	0.68
SBERT	94.5	89.3	91.2	86.9	87.7	86.4	89.33	0.105	0.90
Proposed	95.7	92.4	92.3	90.7	87.5	88.5	91.18	0.068	2.71

Movie Review CLASSIFICATION RESULTS: ZERO-SHOT

Baseline	*en*	*de*	*es*	*ar*	*hi*	*zh*	Avg.	Training Time (hr)	Inference Time (msec)
XLM-R-large	51.2	49.5	49.3	43.7	44.7	47.5	47.66	9.317	4.83
XLM-R-base	50.3	47.0	48.3	41.7	41.4	47.1	45.97	0.367	1.57
mBERT-base	43.8	35.5	39.8	31.6	32.1	40.8	37.27	0.305	1.49
DistilBERT	42.0	34.9	33.8	30.2	29.2	36.7	34.47	0.174	0.76
LASER	43.3	42.4	43.4	42.5	39.4	39.3	41.72	0.011	1.35
SBERT	47.2	44.7	45.5	44.2	43.3	45.1	45.00	0.011	1.89
Proposed	51.7	48.0	49.4	46.2	45.7	49.5	48.42	0.019	5.76

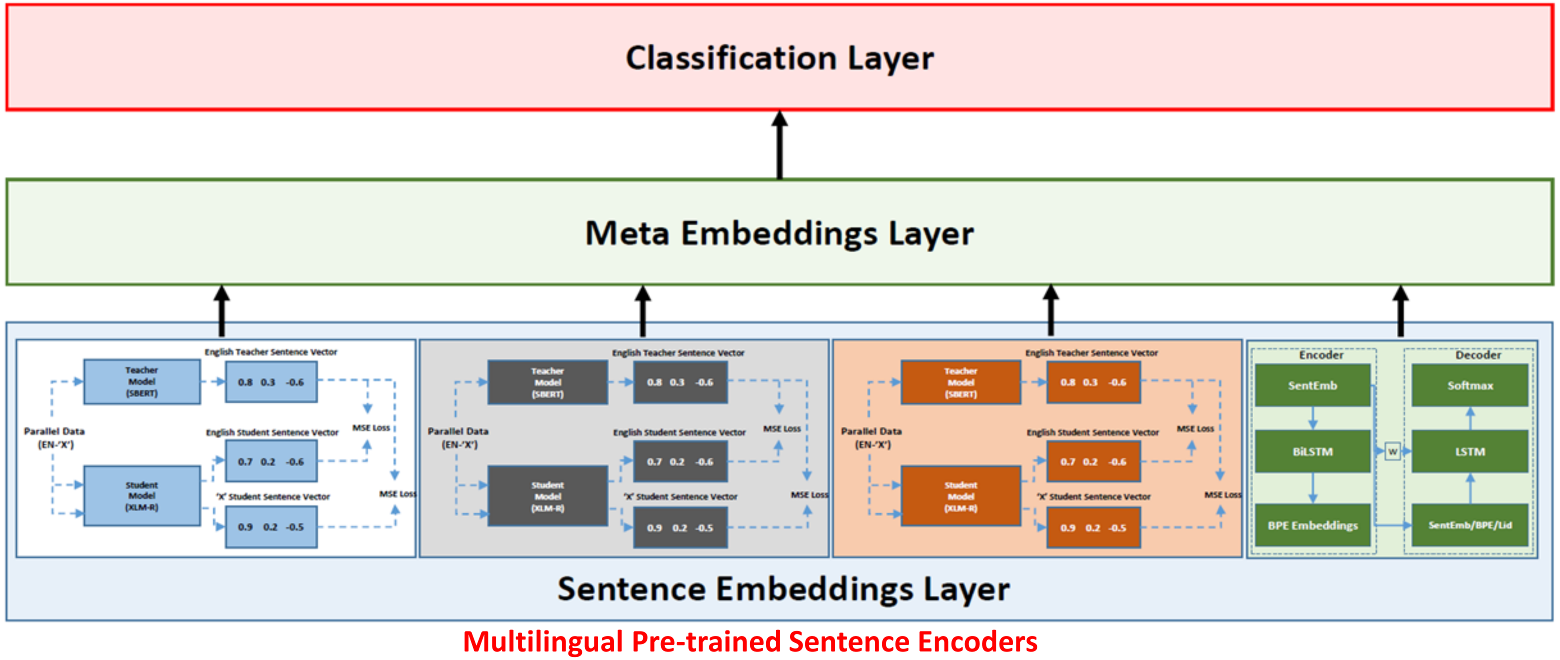
Huawei Service CATEGORIZATION RESULTS: ZERO-SHOT

Baseline	*zh*	*en*	*es*	*ar*	Avg.	Training Time (hr)	Inference Time (msec)
XLM-R-large	68.0	47.7	44.9	47.9	52.12	6.254	3.95
XLM-R-base	80.9	45.1	26.5	38.5	47.75	1.262	1.31
mBERT-base	79.6	13.0	7.0	11.0	27.65	1.660	1.36
DistilBERT	80.5	15.4	6.8	20.6	28.32	0.670	0.73
LASER	73.2	29.1	31.5	31.1	41.2	0.049	1.78
SBERT	69.8	52.5	53.5	57.3	58.2	0.031	2.12
Proposed	75.7	61.5	55.7	60.3	63.3	0.035	5.57

- A multilingual practical and efficient framework for sentence classification.
- Competitive accuracy even outperforming XLM-R-large on several datasets especially for the zero-shot setting.
- Significant gains in training time, with limited training data.

Text Classification – In a nutshell

1. *Significantly low training time*
2. *Requirement of relatively inexpensive compute*
3. *Competitive performance; and*
4. *Robustness across application domains and limited training.*



Efficient Multi-Lingual Sentence Classification Framework with Sentence Meta Encoders

Raj Nath Patel, Edward Burgin, Haytham Assem, Sourav Dutta
Huawei Research Centre, Dublin, Ireland
Email: {raj.nath.patel, edwardburgin, haytham.assem, sourav.dutta2}@huawei.com

Huawei Multi-Lingual E2E Stack (HAMLET)

We have successfully shown the impact of HAMLET light weight engine on various CBG Services:

- **Smart Customer Cloud Service:**
  1. Unsupervised Aspect based Sentiment Analyzer **(70 F1 Score)**
  2. Huawei Account Intent Classification **(75 F1 Score)**
  3. Huawei Petal Search Understanding **(90 F1 Score)**

- **AI Trust & Safety:**
  1. Religious comment Detection **(90 F1 Score)**
  2. Spam Detection **(84 F1 Score)**

- **Huawei Browser:**
  1. Part-of-Speech Tagging **(80 F1 Score)**
  2. News Classification (Ongoing)

Huawei Multi-Lingual E2E Stack (HAMLET)

We have successfully shown the impact of HAMLET light weight engine on various CBG Services:

- **Smart Customer Cloud Service:**
    1. Unsupervised Aspect based Sentiment Analyzer **(70 F1 Score)**
    2. Huawei Account Intent Classification **(75 F1 Score)**
    3. Huawei Petal Search Understanding **(90 F1 Score)**

- **AI Trust & Safety:**
    1. Religious comment Detection **(90 F1 Score)**
    2. Spam Detection **(84 F1 Score)**

- **Huawei Browser:**
    1. Part-of-Speech Tagging **(80 F1 Score)**
    2. News Classification (Ongoing)

THE UNIVERSITY OF EDINBURGH

UNIVERSITY OF LIVERPOOL

NUI Galway OÉ Gaillimh

Trinity College Dublin The University of Dublin

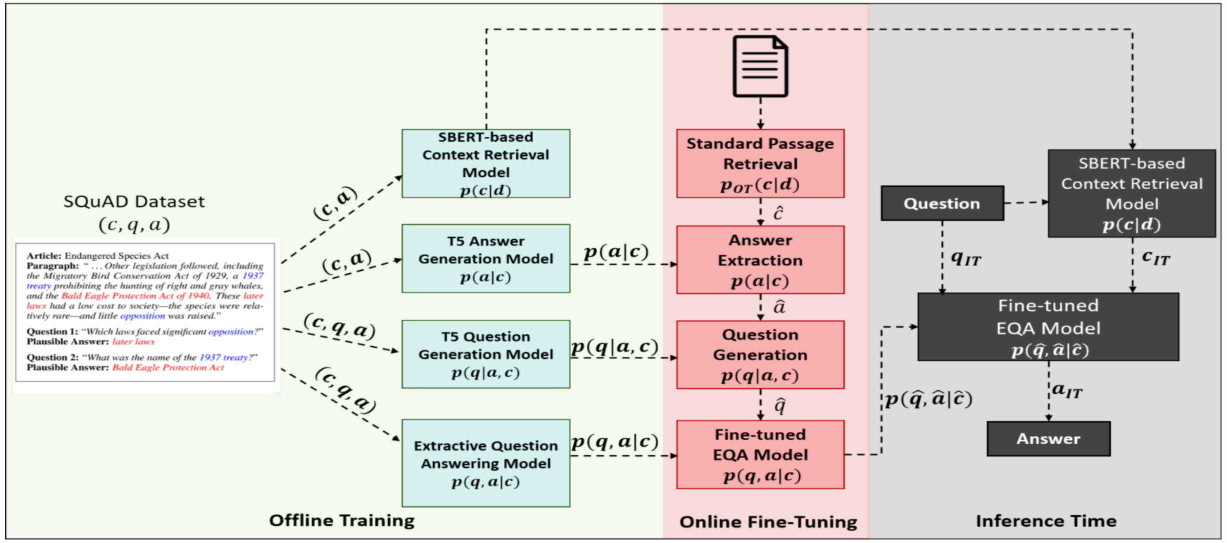
TECHNISCHE UNIVERSITÄT DARMSTADT

Technical & Research Topics

- Multilingual Language Models
- Intent Extraction
- Named Entity Recognition
- Question Answering
- Knowledge Graphs for NLU
- Aspect Based Sentiment Analysis

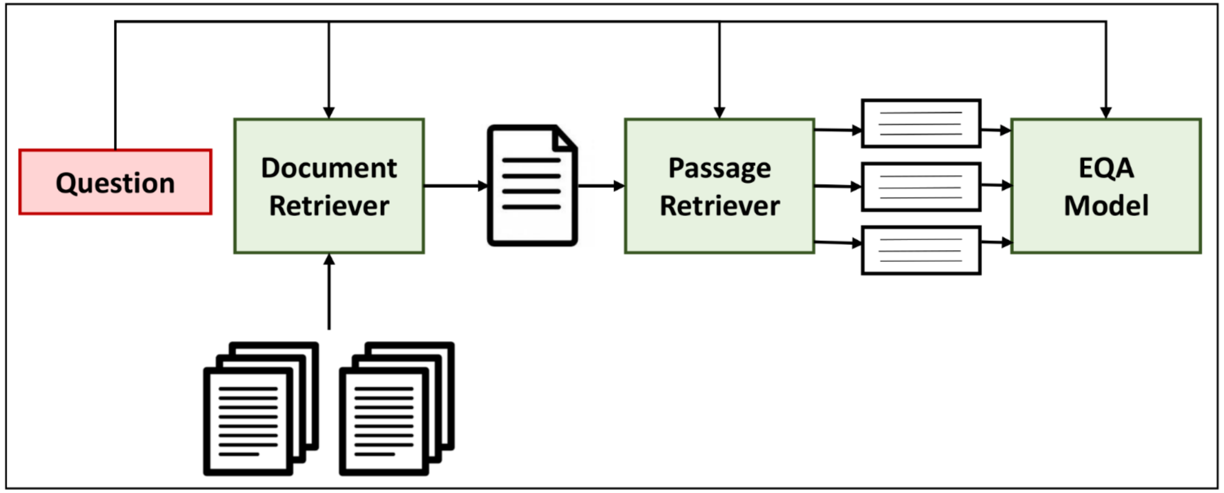
DOMAIN-AWARE QUESTION ANSWERING

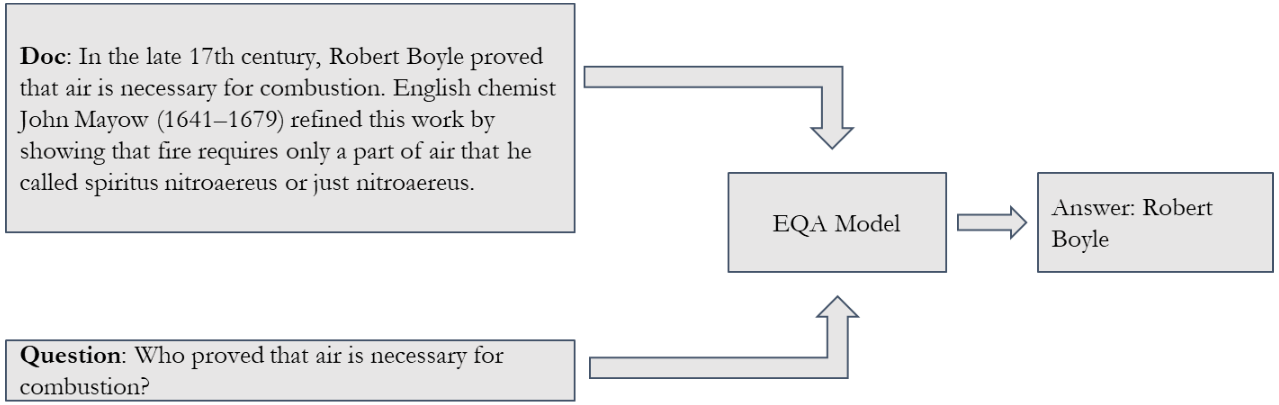
The Automation of User Query Assistance



Extractive Question Answering

- Unbounded user questions.
- Expensive creation of FAQ.
- Precise and targeted answering.
- Use of free-flowing natural language documents.

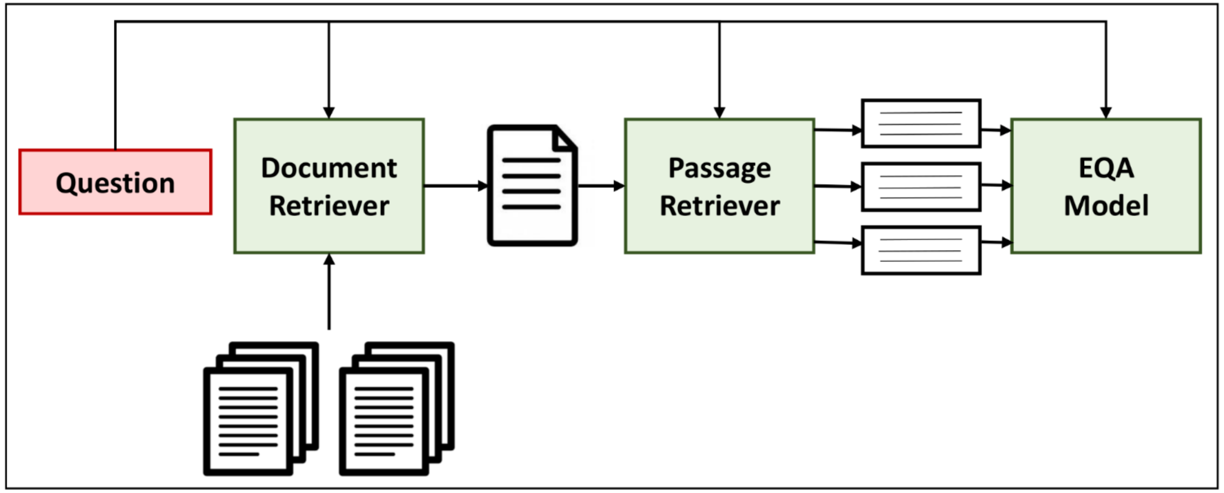




Z. Zhang, H. Zhao, and R. Wang (2020). "Machine Reading Comprehension: The Role of Contextualized Language Models and Beyond". Computational Linguistics.

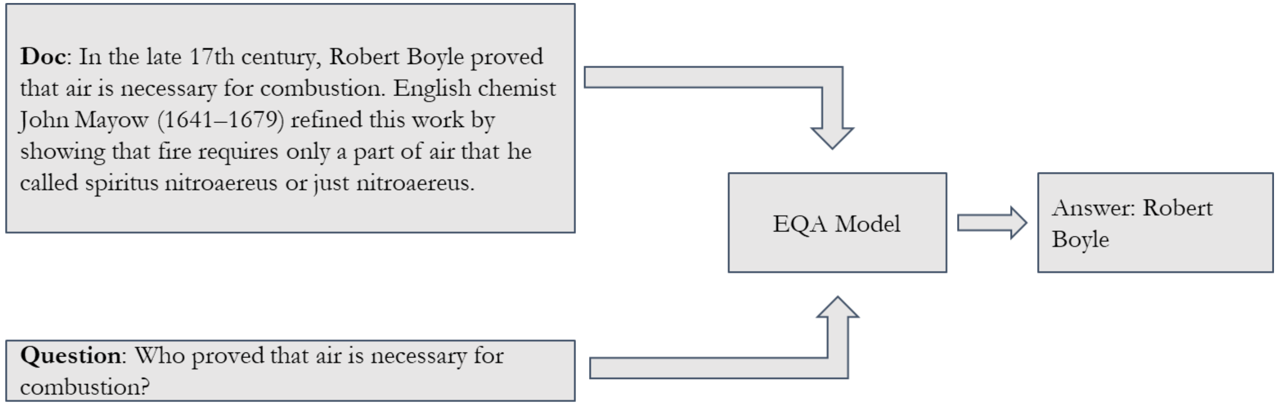
Extractive Question Answering

- Unbounded user questions.
- Expensive creation of FAQ.
- Precise and targeted answering.
- Use of free-flowing natural language documents.



- **<u>Domain Adaptation</u>**: Existing EQA pre-trained perform rather poorly in closed-domain and industrial scenarios.

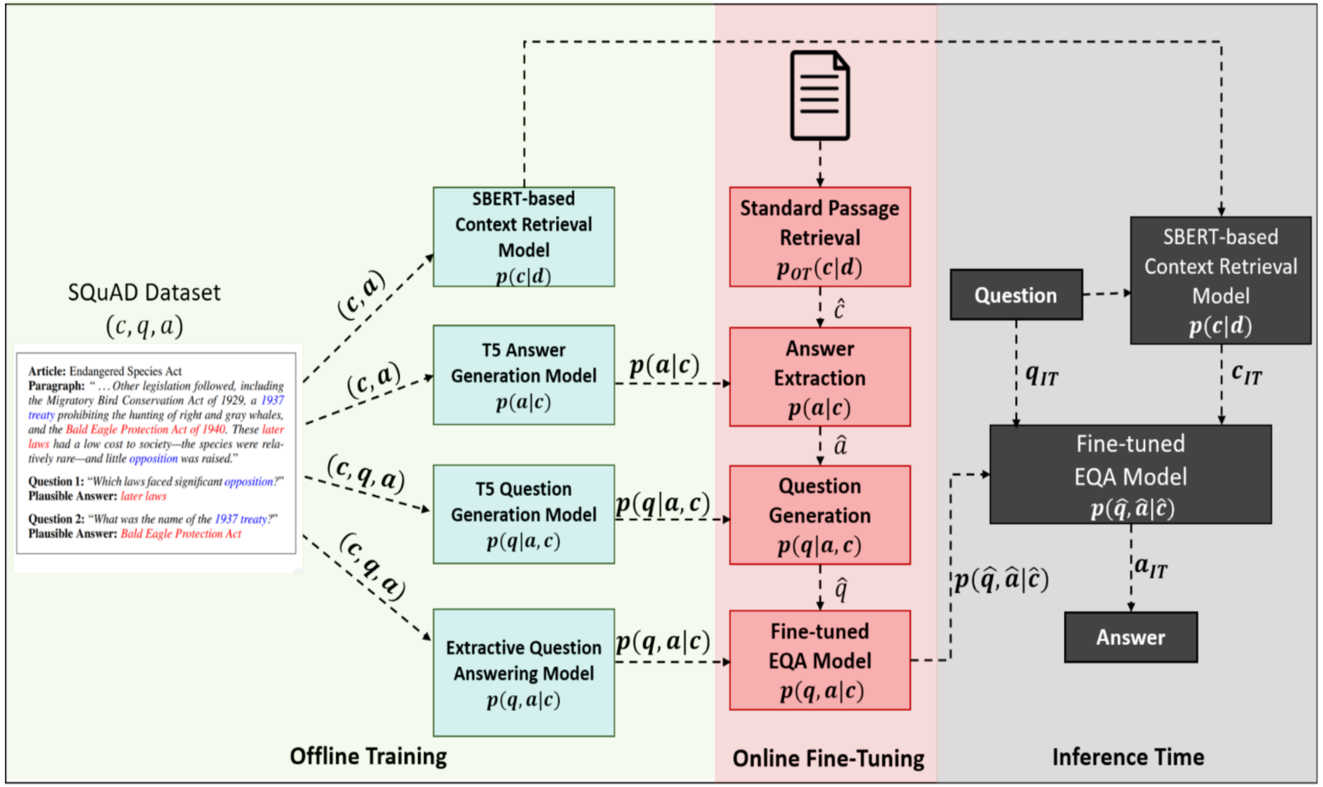
- **<u>Domain-specific Training</u>**: Poor availability and the expensive annotation of domain-specific data.



Z. Zhang, H. Zhao, and R. Wang (2020). "Machine Reading Comprehension: The Role of Contextualized Language Models and Beyond". Computational Linguistics.

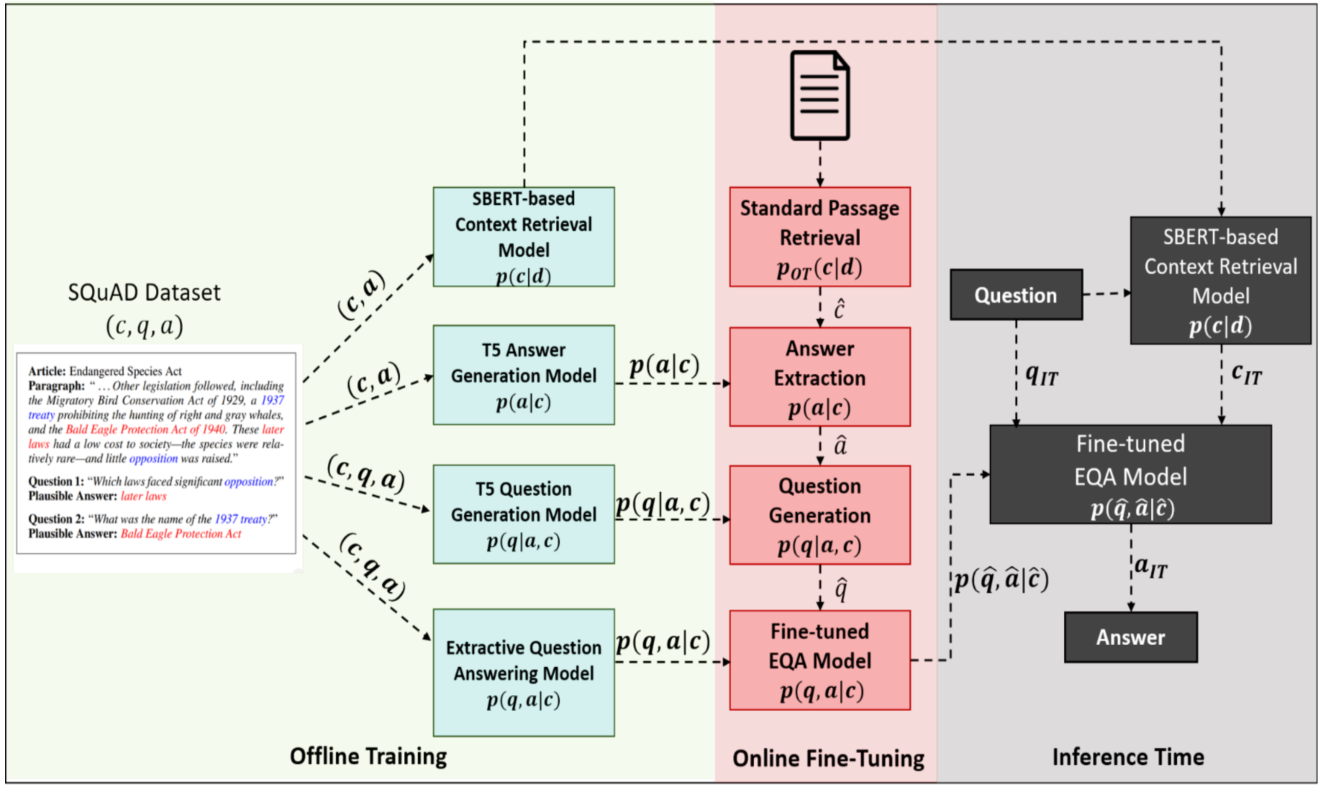
Self-Learning for QA

- *Self-learning domain adaptation:*
  - <Question, Answer, Context> triples generation using T5 model.
  - SBERT tuned to identify relevant sentences as NLI formulation.
  - QA model fine-tuned on triples using synthetic domain data.



Self-Learning for QA

- *Self-learning domain adaptation:*
  - <Question, Answer, Context> triples generation using T5 model.
  - SBERT tuned to identify relevant sentences as NLI formulation.
  - QA model fine-tuned on triples using synthetic domain data.



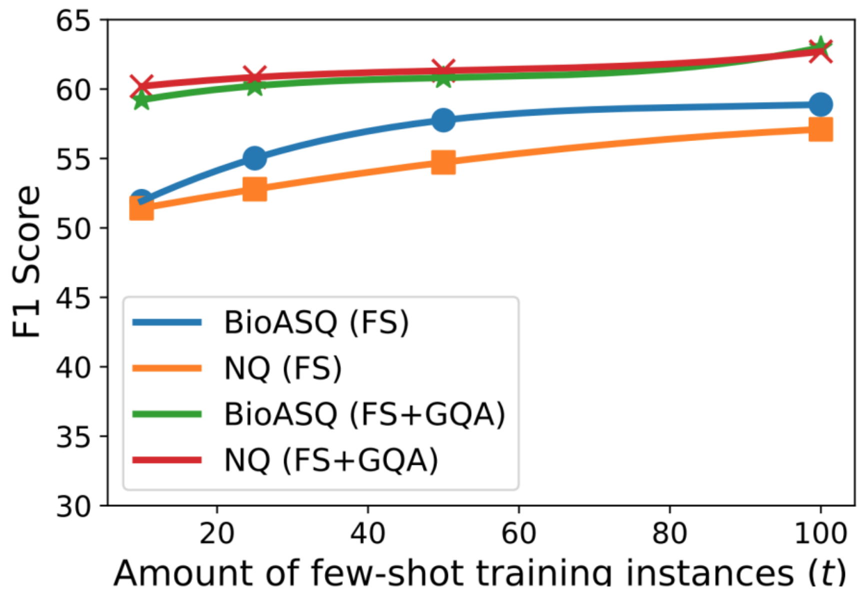
- On arrival of user question:
  - Prune text to identify candidate context.
  - Pass pruned context to EQA model.

Evaluating Domain-Based QA

Model	NQ		DROP		NewsQA		BioASQ		TextBookQA	
	EM	F1	EM	F1	EM	F1	EM	F1	EM	F1
SpanBERT-SQuAD	35.21	49.71	24.08	29.91	37.93	53.19	38.27	49.04	28.94	37.00
SpanBERT-SQuAD-GQA‡	**45.18**	**60.95**	31.71	40.99	**40.42**	**57.40**	46.94	58.88	33.46	44.81
QASAR‡	44.18	59.76	**32.67**	**41.06**	40.21	56.63	**52.21**	**63.70**	**35.32**	**47.02**

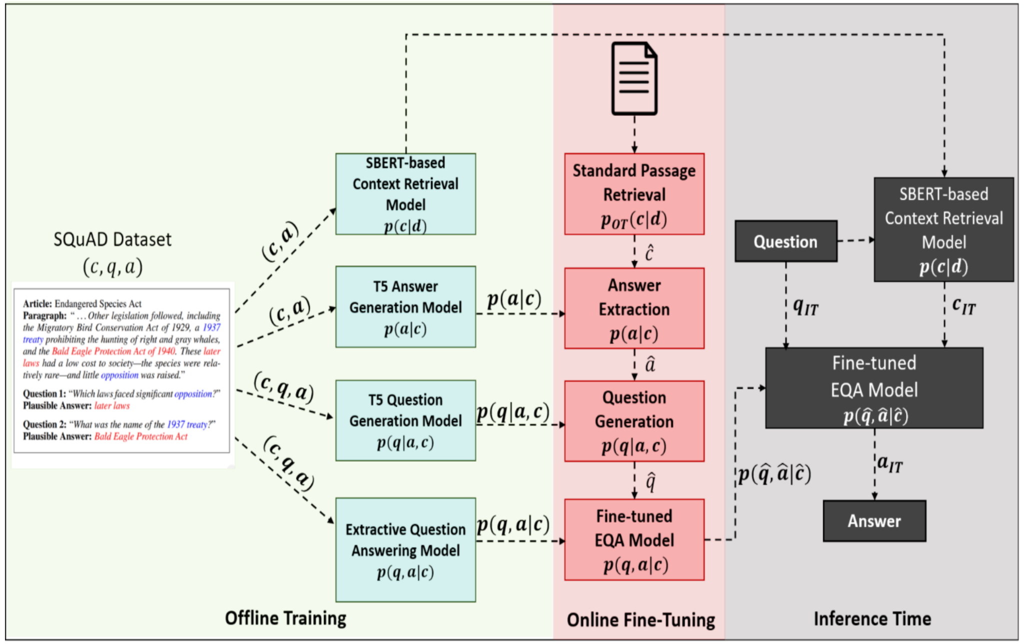
‡ Results are *statistically significant* with $p < 1e-5$ based on paired t-test with *SpanBERT-SQuAD*.

- We observe significant performance improvements in understanding domain-specific text.
  - Applicable also for open-domain QA solutions.
- Quality of user-experience is better than industrial solutions.
- Few-shot training further boosts accuracy.



Domain-Based QA – In a nutshell

- A novel and efficient question answering framework for self-learning based domain adaptation on closed-domain applications.
- Showcased the use of Language Models for eliminating the need of annotated domain-specific training dataset.
- Performance better than existing methodologies.



Semantic Aware Answer Sentence Selection using Self-Learning based Domain Adaptation

Rajdeep Sarkar*
Data Science Institute, National University of Ireland Galway
Galway, Ireland
rajdeep.sarkar@insight-centre.org

Sourav Dutta
Huawei Research
Dublin, Ireland
sourav.dutta2@huawei.com

Haytham Assem†
Amazon Alexa AI
Cambridge, UK
hithsala@amazon.co.uk

Mihael Arcan
Data Science Institute, National University of Ireland Galway
Galway, Ireland
mihael.arcan@insight-centre.org

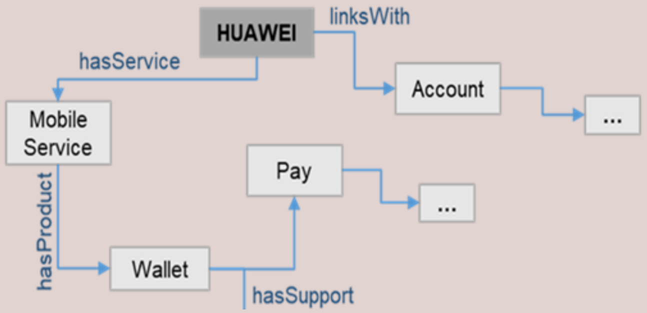
John McCrae
Data Science Institute, National University of Ireland Galway
Galway, Ireland
john.mccrae@insight-centre.org

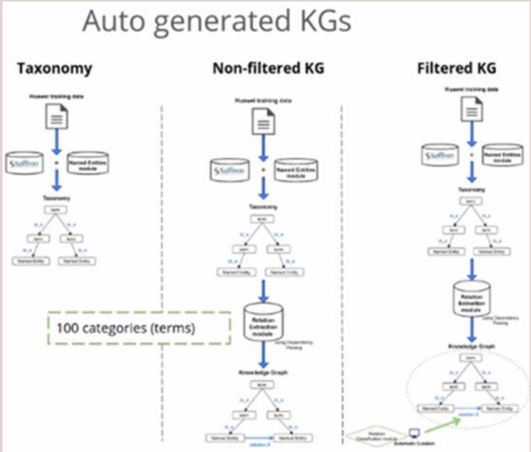
QASAR: Self-Supervised Learning Framework for Extractive Question Answering

Haytham Assem*
Huawei Research
Dublin, Ireland
haytham.assem@huawei.com

Rajdeep Sarkar*†
National University of Ireland Galway
Galway, Ireland
r.sarkar1@nuigalway.ie

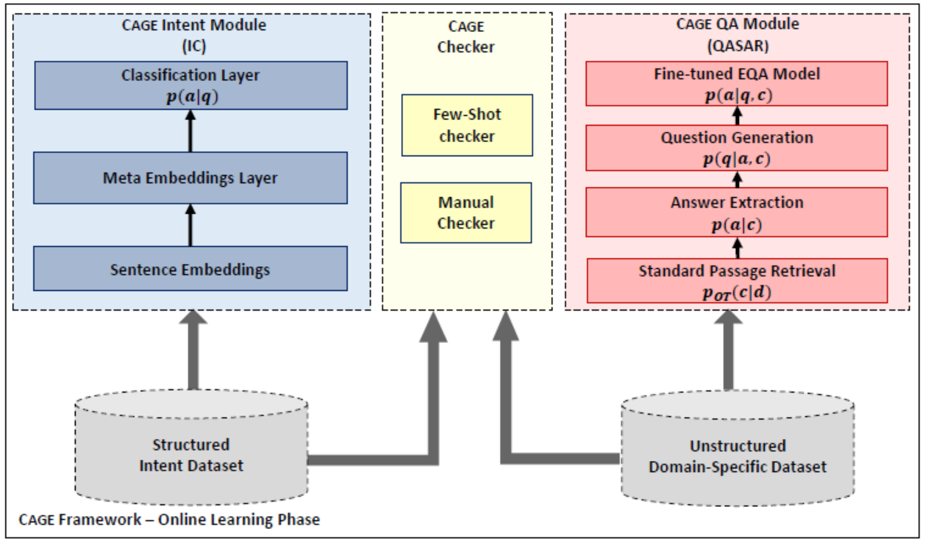
Sourav Dutta
Huawei Research
Dublin, Ireland
sourav.dutta2@huawei.com

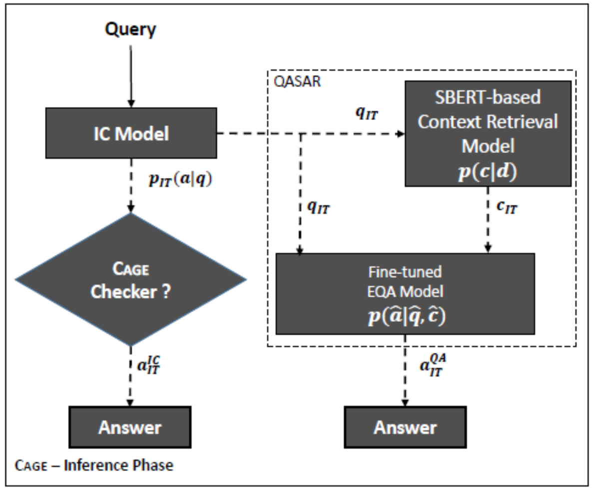




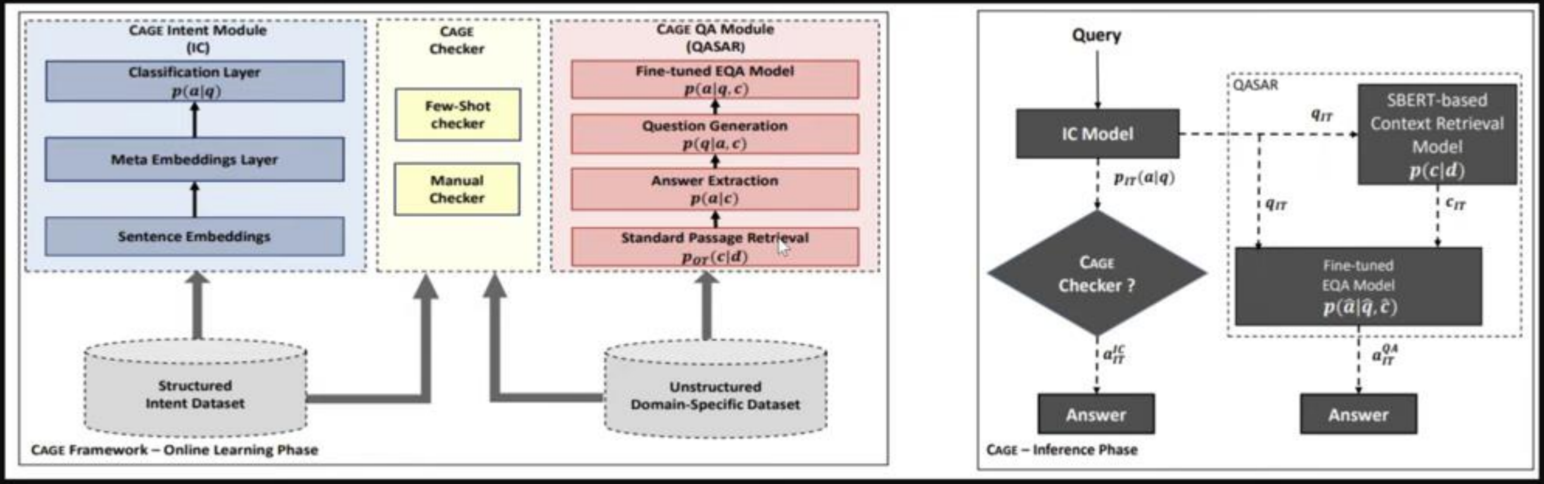
ONE-STOP INFORMATION HUB

The Intelligent Chatbot Solution





The Intelligent Chatbot Solution

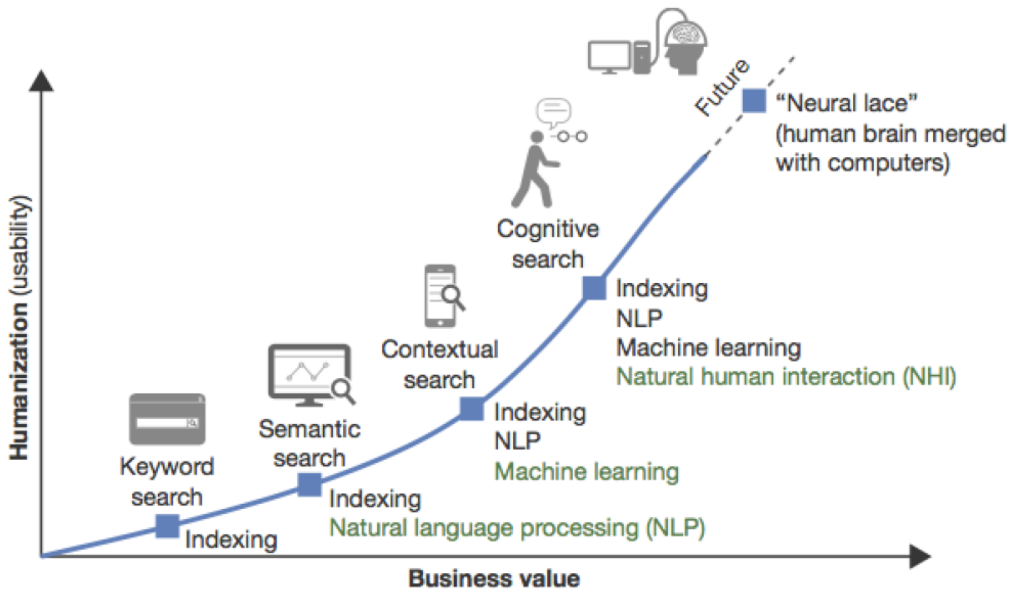


Research Contributions

Domain	Paper Title	Conference
Text Embedding	*Enhanced Sentence Meta-Embeddings for Textual Understanding*	**ECIR 2022**
	Cross-lingual Sentence Embedding using Multi-Task Learning	**EMNLP 2021**
	Sequence-to-Sequence Learning on Keywords for Efficient FAQ Retrieval	**IJCAI 2021**
	DTAFA: Decoupling Training Architecture for FAQ Retrieval	**SIGDIAL 2021**
	Efficient Multi-Lingual Sentence Classification Framework with Sentence Meta Encoders	**IEEE-BigData 2021**
	MUFIN: Enriching Semantic Understanding of Sentence Embedding using Dual Tune Framework	**IEEE-BigData 2021**
Question - Answering	*Semantic Aware Answer Sentence Selection using Self-Learning based Domain Adaptation*	**KDD 2022**
	CAGE: A Hybrid Framework for Closed-Domain Conversational Agents	**ECML-PKDD 2022**
	QASAR: Self-Supervised Learning Framework for Extractive Question Answering	**IEEE-BigData 2021**
Network Pruning	*Aligned Weight Regularizers for Pruning Pretrained Neural Networks*	**ACL 2022**
	Self-Distilled Pruning of Deep Neural Networks	**ECML-PKDD 2022**

Research Contributions

Domain	Paper Title	Conference
Text Embedding	*Enhanced Sentence Meta-Embeddings for Textual Understanding*	**ECIR 2022**
	Cross-lingual Sentence Embedding using Multi-Task Learning	**EMNLP 2021**
	Sequence-to-Sequence Learning on Keywords for Efficient FAQ Retrieval	**IJCAI 2021**
	DTAFA: Decoupling Training Architecture for FAQ Retrieval	**SIGDIAL 2021**
	Efficient Multi-Lingual Sentence Classification Framework with Sentence Meta Encoders	**IEEE-BigData 2021**
	MUFIN: Enriching Semantic Understanding of Sentence Embedding using Dual Tune Framework	**IEEE-BigData 2021**
Question - Answering	*Semantic Aware Answer Sentence Selection using Self-Learning based Domain Adaptation*	**KDD 2022**
	CAGE: A Hybrid Framework for Closed-Domain Conversational Agents	**ECML-PKDD 2022**
	QASAR: Self-Supervised Learning Framework for Extractive Question Answering	**IEEE-BigData 2021**
Network Pruning	*Aligned Weight Regularizers for Pruning Pretrained Neural Networks*	**ACL 2022**
	Self-Distilled Pruning of Deep Neural Networks	**ECML-PKDD 2022**



JOIN US IN

BUILDING A BETTER CONNECTED WORLD

THANK YOU