

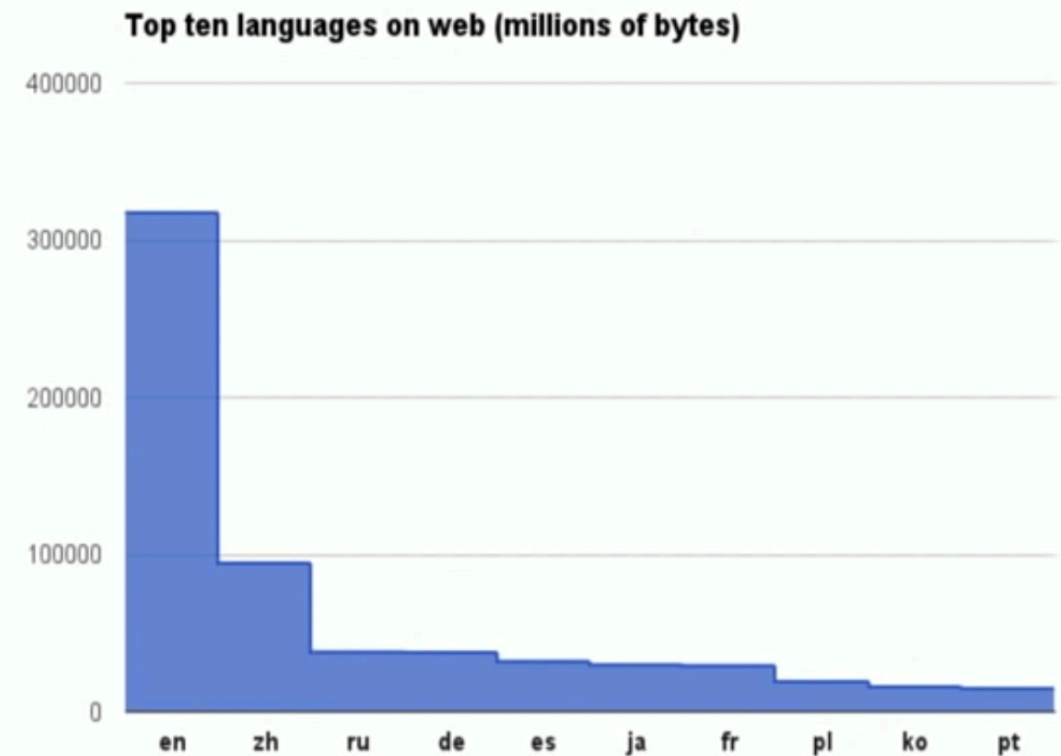


NEURAL MACHINE TRANSLATION

- Aiswarya Ramachandran

WHY IS MACHINE TRANSLATION IMPORTANT?

- **50%** of Internet content is in English.
- Only **20%** of the world's population speaks English.



ARE TRANSLATIONS ONLY FOR COMPANIES LIKE GOOGLE?

WHAT IS SEQUENCE TO SEQUENCE MODELLING?

Machine Language Translation

*Les modèles de séquence
sont super puissants*

Sequence Model

*Sequence models are super
powerful*

Text Summarization

*A strong analyst have 6
main characteristics. One
should master all 6 to be
successful in the industry :
1.
2.*

Sequence Model

*6 characteristics of
successful analyst*

Chatbot

How are you doing today?

Sequence Model

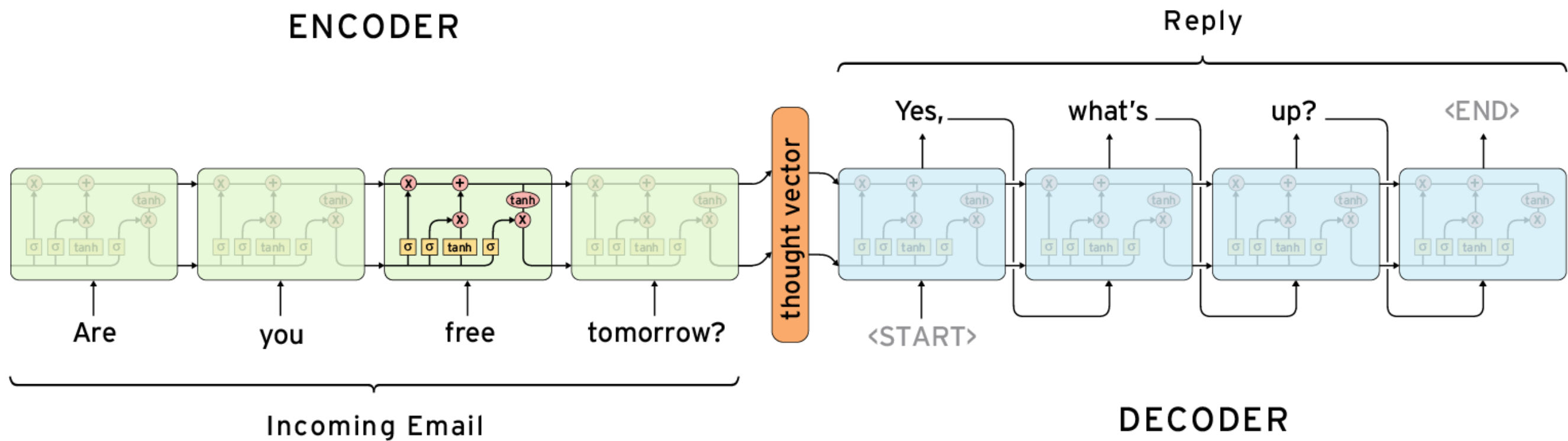
*I am doing well. Thank you.
How are you doing today?*

Source <https://www.analyticsvidhya.com/blog/2018/04/sequence-modelling-an-introduction-with-practical-use-cases/>

Problem Statement

- **Translate Text from English to Hindi using Seq2Seq Modelling**
- **We are using Word Level Translation**
- **The data for this problem is available on Kaggle**
- **The methodology used here can be extended to any sequence to sequence problem**

ENCODER-DECODER ARCHITECTURE



TEACHER FORCING DURING TRAINING

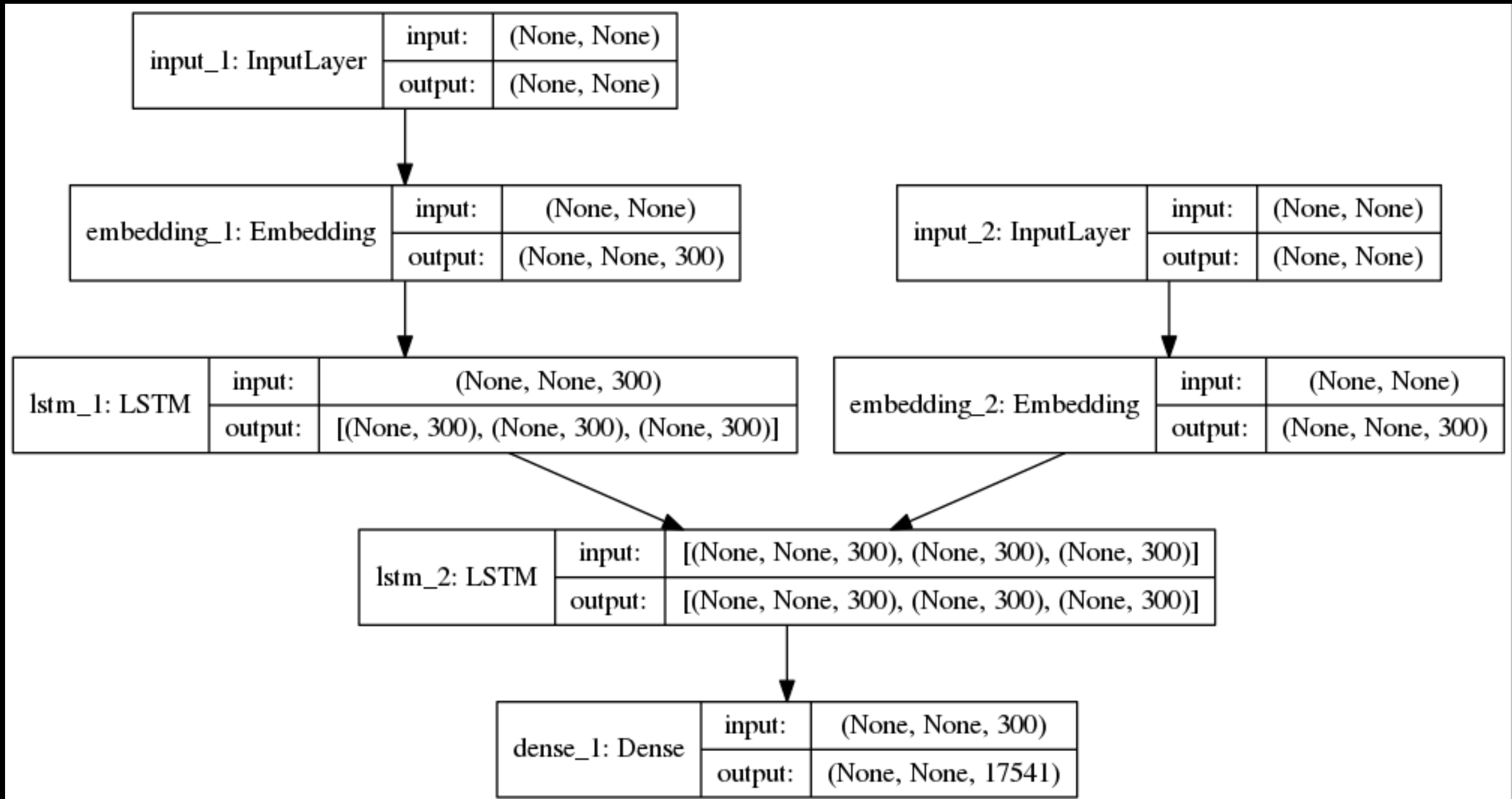
- Decoder behaves differently in Training and Inference
- During training, we use a mechanism called Teacher Forcing
 - “Teacher Forcing is a mechanism in which during the training the model receives the ground truth output $y(t)$ as input at time $t+1$ ”

Consider the following example:

मैं बैंगलोर में रहता हूँ

NAIVE APPROACH			TEACHER FORCING	
X	yhat		X	yhat
START	हम		START	हम
START,हम	बैंगलोर		START,मैं	बैंगलोर
START,हम,बैंगलोर	रहते		START,मैं,बैंगलोर	रहते
START,हम,बैंगलोर,रहते	हैं		START,मैं,बैंगलोर,में	रहता
START,हम,बैंगलोर,रहते,हैं	END		START,मैं,बैंगलोर,में,रहता	हैं
			START,मैं,बैंगलोर,में,रहता,हूँ	END

ARCHITECTURE OF THE MODEL



CONCLUSION

- **Add more data**
- **Attention Mechanisms can be used to reduce the loss**
- **Use Dropouts and other regularization techniques**
- **Use Bidirectional LSTMs in the Encoder or multilayered LSTMs**
- **Use Beam Search**
- **Use BLEU Score to evaluate the model**