

Specialist Diploma in Artificial Intelligence

AI Applications with Deep Learning

Chatbot

Sub topics

- **What is a Chatbot?**
- **Traditional Chatbot Intent Recognition**
- **Traditional Chatbot Custom Named Entity Recognition**
- **LLM Chatbot**
- **Traditional Chatbot vs LLM-based Chatbot**

Creating Your Own Chatbot

WHAT IS A CHATBOT?

Open Domain Chatbot

A multi-turn chatbot handles conversations with multiple exchanges, maintaining context throughout the interaction. (e.g. LLM Chatbot)

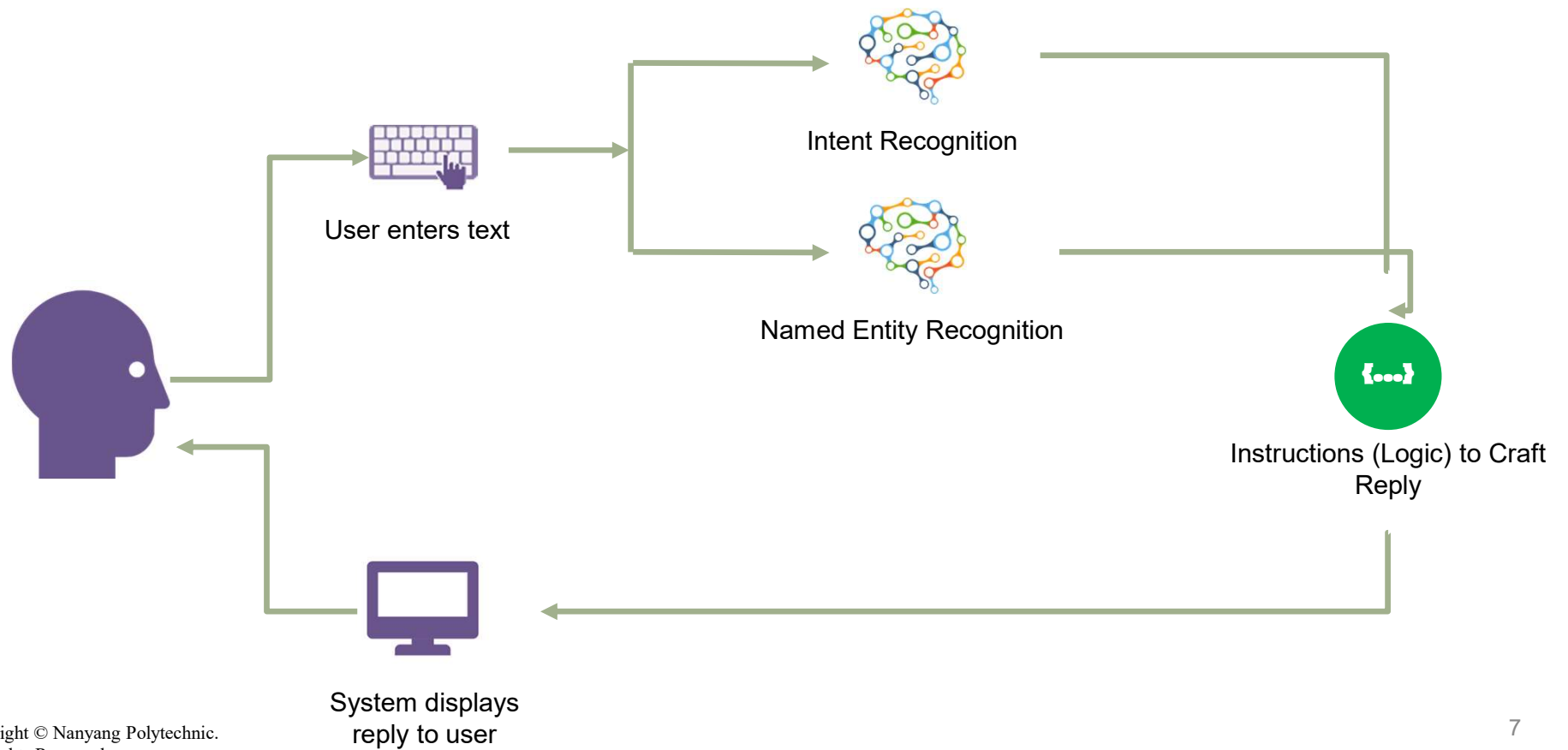
- **Context Retention:** Remembers previous inputs and references earlier parts of the conversation.
- **Coherent Responses:** Provides relevant and logical answers based on the entire conversation.
- **Complex Interactions:** Handles tasks that require multiple steps or follow-up questions (e.g., booking, troubleshooting).
- **State Management:** Tracks the conversation flow and adapts based on user input.

A Simple Chatbot – Closed Domain

- What we will try to build:
 - Closed Domain = Chatbot meets specific business objectives
- For each text said by the user:
 - Use AI to Classify Intent (Intent Classification)
 - Use AI to Extract Entities (Entity Recognition)
 - Write Rules to Process Intent and Entities
 - Write Code to Display Reply to User

See pra

Chatbot

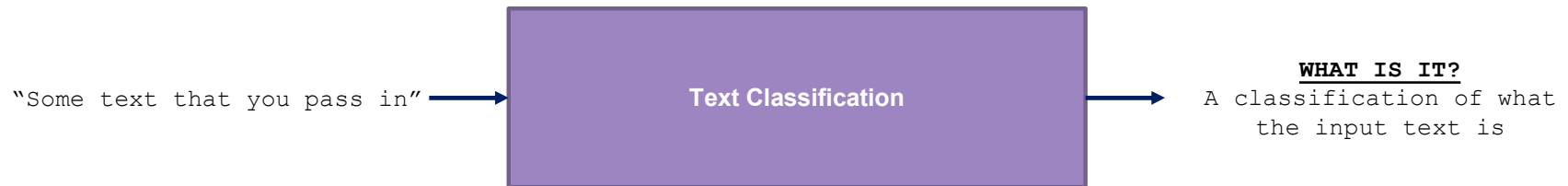


Creating Your Own Chatbot

INTENT RECOGNITION – ZERO SHOT CLASSIFICATION

Text Classification

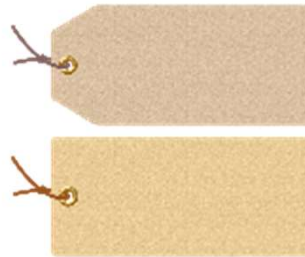
- Intent recognition can be achieved with “Sequence / Text Classification”.
- What we are actually doing:



Some Use Cases



Social Media Monitoring
(Complaints, Threats)



Automated Tagging of
User Content
(Tag Suggestions, Categorization)



Support Ticket Classification /
Intent Recognition



Document Classification
and Search



Sentiment Analysis
(Branding, Public Relations)

Two Ways to Do This

- Zero-Shot Text Classification

Zero-shot means:

Use a pre-trained model with no *additional* training data for any generic task (this task was never seen in the original training data set).

This is good if you have no training data. Less accurate than training with your own custom dataset

- Collect Data and Perform Transfer Learning on the Pre-Trained Model with your Custom Labels

Transformers with HuggingFace – Zero-Shot



"Can I know what flights are
available?"

"**flight**"-related?

Zero-Shot Classification Pipeline
(Hugging Face)

0.9

"Can I know what flights are
available?"

"**greeting**"-related?

Zero-Shot Classification Pipeline
(Hugging Face)

0.1

Highest score wins

With Zero-Shot Learning

- The candidate “label” are English labels, and the model tries to compute the similarity the input text’s meaning with the label’s meaning.
- So to achieve good results, the choice of the candidate labels (with clear semantic boundary with the rest of the labels) is very important.

Creating Your Own Chatbot

INTENT RECOGNITION – TRAINING WITH CUSTOM LABELS

Dataset

- We want to build a chatbot that interacts with customers on **flight enquiries** and other flight-related matters.
- Assuming that we use a combination of the following datasets:
 - <https://www.kaggle.com/hassanamin/atis-airlinetravelinformationsystem>
ATIS that captured customers' requests and classified each of them
 - <https://github.com/alyssaong1/botframework-smalltalk/blob/master/smalltalkkb.tsv>
A small talk chatbot dataset with greetings, yes/no answers and other small talk

Transformers with HuggingFace

Custom Trained Labels

"Can I know what flights are available?"



Transformer Model
(Hugging Face)



chat_abbreviation	: 0.0
chat_flight	: 0.8
chat_yes	: 0.1
others	: 0.1

This is what we want to achieve:

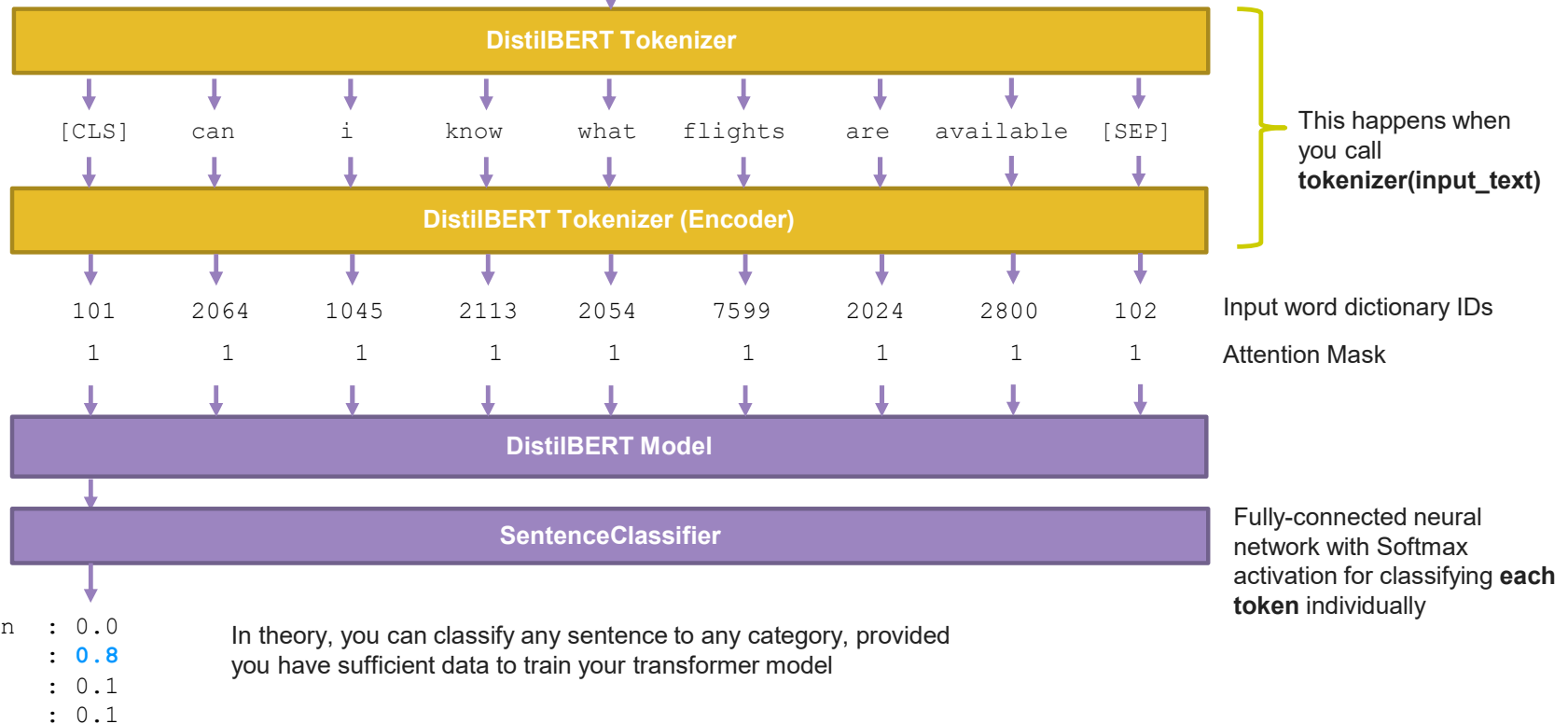
- Pass in a sentence as input,
- Retrieve a classification as output

Transformers with HuggingFace

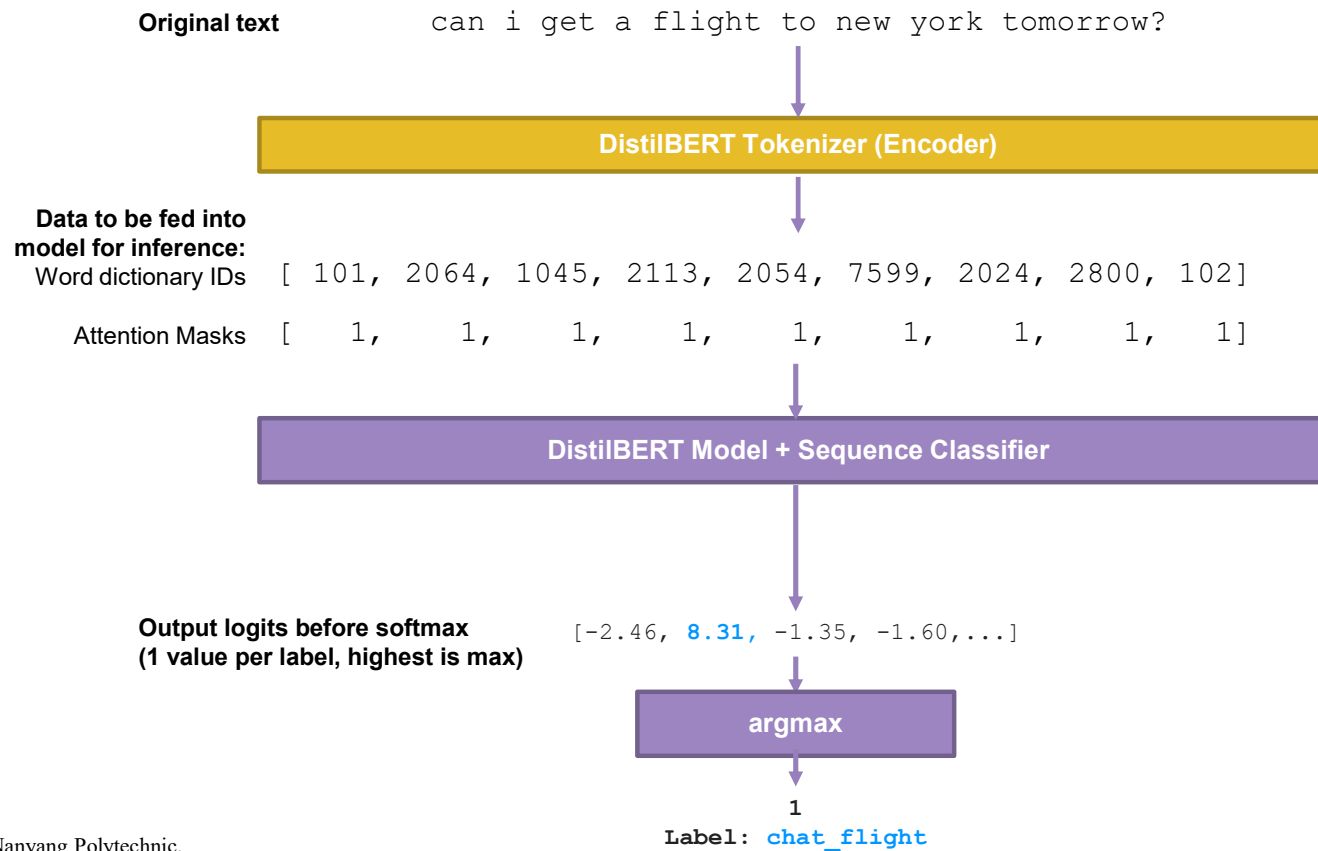
Custom Trained Labels



"Can I know what flights are available?"



Inference



Creating Your Own Chatbot

CUSTOM NAMED ENTITY RECOGNITION

What is Named Entity Recognition?

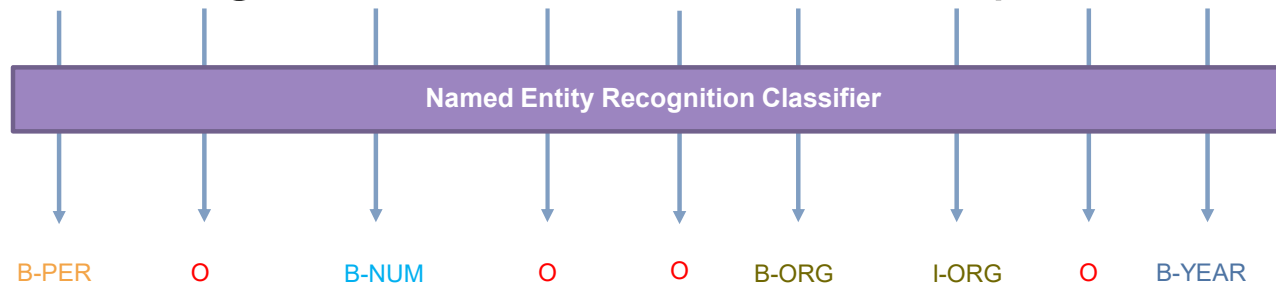
Seeks to locate and classify named entity mentions in unstructured text into pre-defined categories such as:

the person names, organizations, locations, medical codes, time expressions, quantities, monetary values, percentages, etc.

https://en.wikipedia.org/wiki/Named-entity_recognition

NER – Used for Extracting Actionable Data

Jim bought 300 shares of Acme Corp. in 2006.



For multi-word entities like “Acme Corp”, we use:

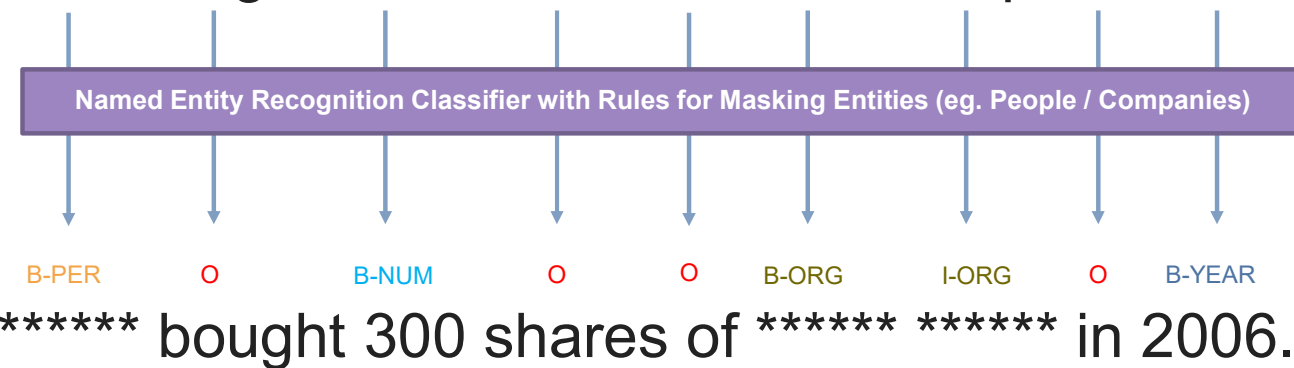
B-ORG (1st word), and

I-ORG (subsequent words joined to the first name)

This helps to extract important information that your app can act upon.

NER – Security and Personal Data Protection

Jim bought 300 shares of Acme Corp. in 2006.



Your app can also use the classified labels for each word to selectively mask them.

The I- O- B- Format

- As mentioned, to label our entities, we use IOB format.
- Also known as **Inside-Outside-Beginning**:

Jim
bought
2
tickets
to
New
York

B-PER

O

O

O

O

B-LOC

I-LOC

This single word "Jim" forms the name of a person

The two words, "New York", form an instance of the location.

The first word of an entity is marked with B-###, second word onwards marked with I-###

Training and Test Dataset

- This is an example of the CoNLL 2003 dataset:
 - <https://github.com/synalp/NER/blob/master/corpus/CoNLL-2003/eng.testa>

-DOCSTART- -X- 0 0

CRICKET NNP I-NP 0

- : 0 0

LEICESTERSHIRE NNP I-NP I-ORG

TAKE NNP I-NP 0

OVER IN I-PP 0

AT NNP I-NP 0

TOP NNP I-NP 0

AFTER NNP I-NP 0

INNINGS NNP I-NP 0

VICTORY NN I-NP 0

. . 0 0

LONDON NNP I-NP I-LOC

1996-08-30 CD I-NP 0

5 Entity Types:

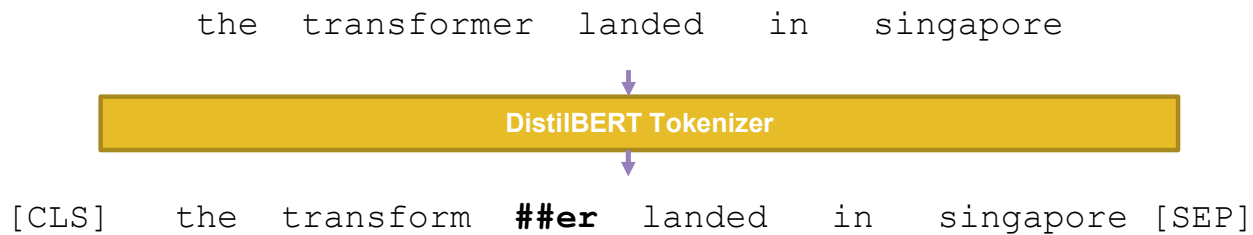
- B-PER / I-PER Person Name
- B-LOC / I-LOC Country or Location Name
- B-ORG / I-ORG Organization or Company Name
- B-MISC / I-MISC Any Other Names
- O Other Words to be Ignored

Handling Sub-words

- The DistilBERT Transformers is trained on a dictionary of about 30,000+ words
- Not all words can exist in the dictionary, some words are represented with multiple tokens of sub-words.
- For example, the word
 - “transformer” is tokenized to 2 words:
 - “transform” + “##er”

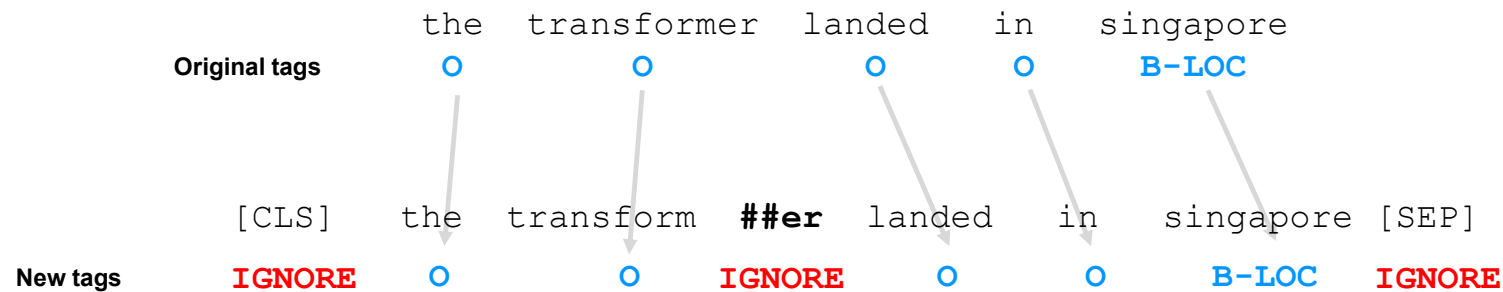
Handling Sub-words

- Big idea:
 - ignore the token labelling and classification for the sub-words.



Handling Sub-words

- Big idea:
 - ignore the token labelling and classification for the sub-words.



Creating Your Own Chatbot

LLM-based Chatbot

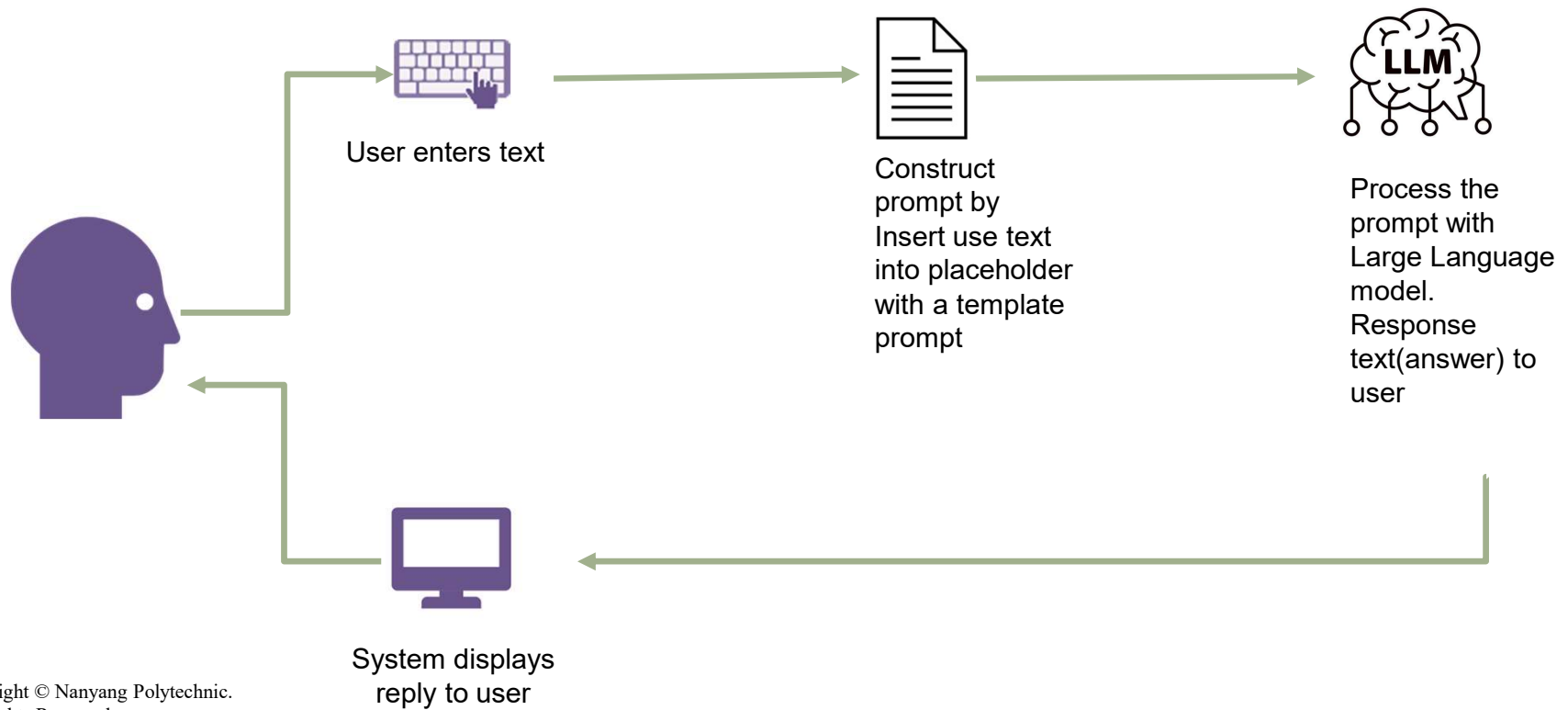
LLM Chatbot

A chatbot powered by Large Language Models (LLMs), advanced AI systems trained on vast text datasets to understand and generate human-like language.

Key Features

- Natural Language Understanding: Handles complex, nuanced queries.
- Context Awareness: Maintains and remembers conversation context.
- Adaptability: Easily handles diverse topics and industries.
- Personalization: Provides tailored responses based on user input.

LLM Chatbot



Creating Your Own Chatbot

Traditional Chatbot vs LLM-based Chatbot

Traditional Chatbot features and limitations

- Key Features:
 - Rule-based systems with decision trees and simpler machine learning algorithms.
 - Operate on predefined scripts, ensuring deterministic and predictable responses.
- Processing Style:
 - Rely heavily on predefined rules for user queries.
 - High consistency but limited flexibility for unexpected queries.
- Limitations:
 - Inflexible to new contexts or scenarios.
 - Minimal learning potential, limiting evolution over time. Struggles to adapt or scale for broader use cases.

LLM based Chatbots

- Advanced Processing Abilities
 - Built on deep learning and transformer architectures (e.g., GPT-4, Claude).
 - Handle intricate language patterns for nuanced communication.
 - Superior capability in processing complex inputs.
- Contextual Awareness
 - Maintain conversation context across multiple interactions.
 - Deliver dynamic, personalised responses.
 - Minimise repetition for a seamless user experience.
- Training and Scalability
 - Extensive training datasets enable adaptability to various topics.
 - Meet diverse industry needs with minimal manual updates.
 - Low maintenance costs and high efficiency.

Summary

- **What is a Chatbot?**
- **Traditional Chatbot Intent Recognition**
- **Traditional Chatbot Custom Named Entity Recognition**
- **LLM Chatbot**
- **Traditional Chatbot vs LLM-based Chatbot**