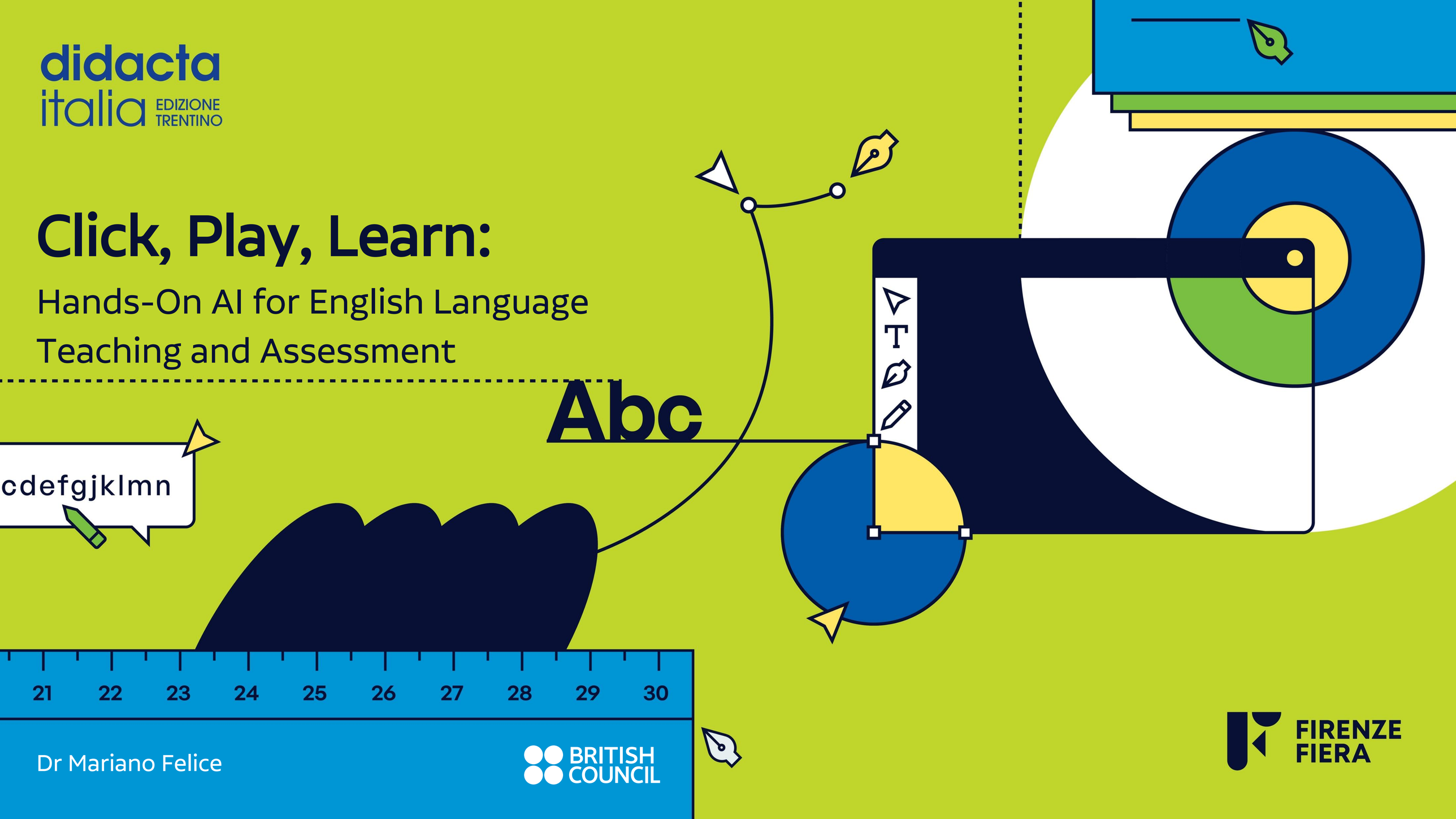
didacta italia EDIZIONE TRENTINO

Click, Play, Learn:

Hands-On AI for English Language Teaching and Assessment



Dr Mariano Felice

BRITISH COUNCIL

FIRENZE FIERA

AGENDA

- Introduction to AI
- Machine Learning
- Language Models
- Generative AI
- Automated Essay Scoring
- Automatic Speech Recognition
- Speech Synthesis (text to speech)

BEFORE WE BEGIN

bit.ly/bc-didacta-2025

WHAT IS AI?

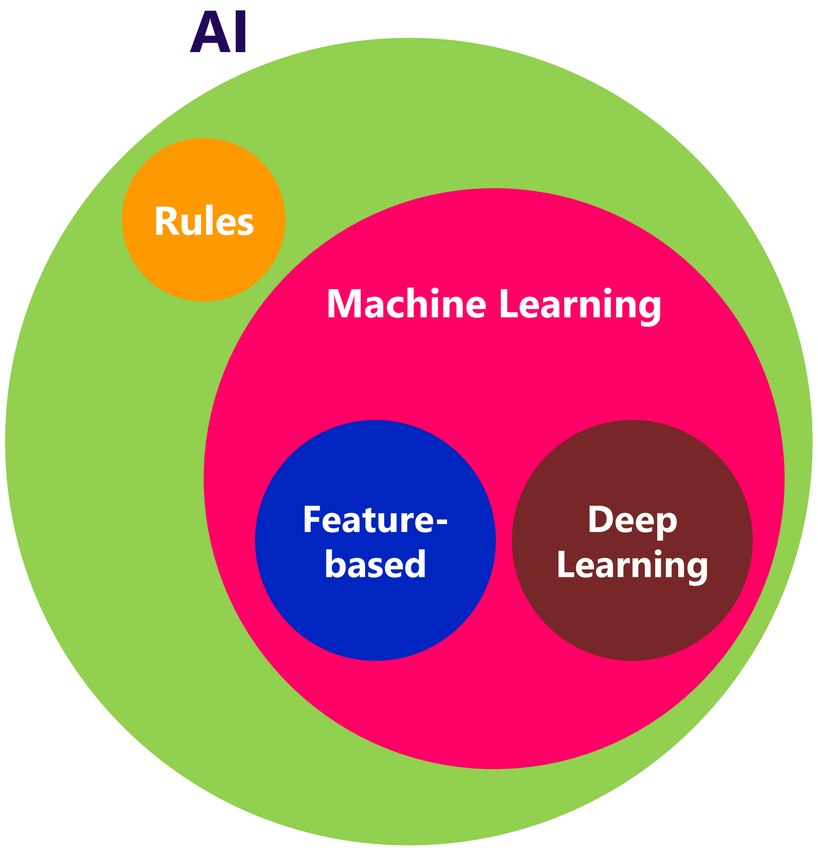
Artificial intelligence is the ability of computers to perform tasks that are typically associated with human intelligence.

AI IS NOT CHATGPT

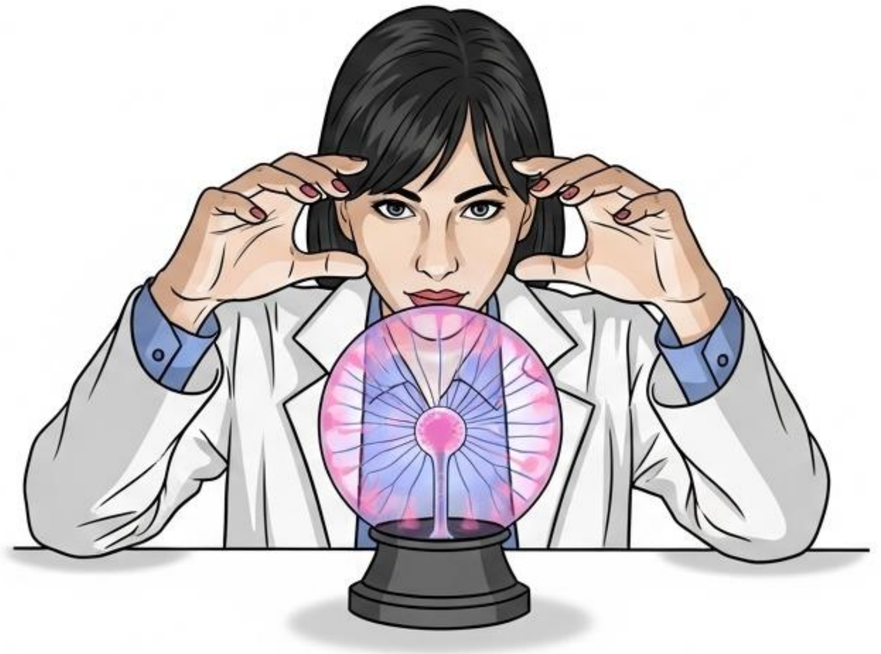
CHATGPT IS AI

AI IS NOT ONLY AUTO-SCORING

HOW DO WE CREATE IT?



The goal of machine learning is to make **predictions**…



The goal of machine learning is to make **predictions**…

using mathematical models.

EVERYTHING IS A PREDICTION!

CLASSIFICATION

- A CEFR level
- Pass/fail
- Language identification

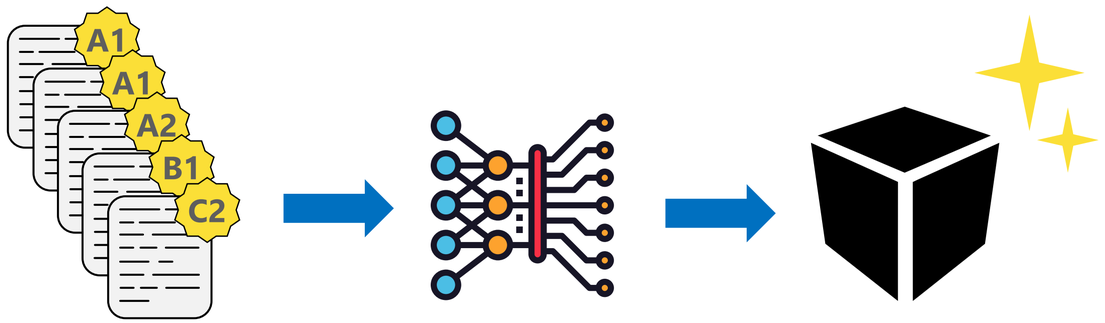
REGRESSION

- A score
- Item difficulty
- Time to proficiency

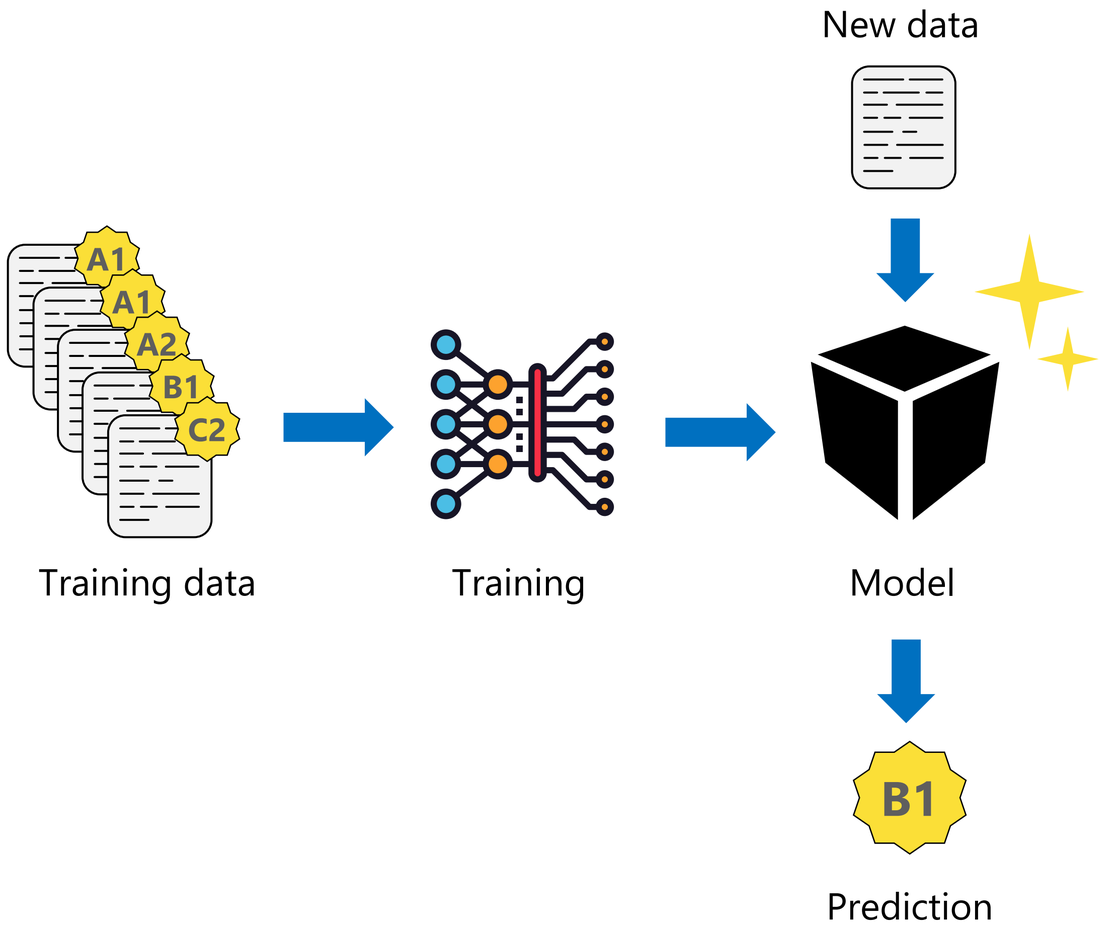
SEQUENCES

- Translation
- Text generation
- Image generation

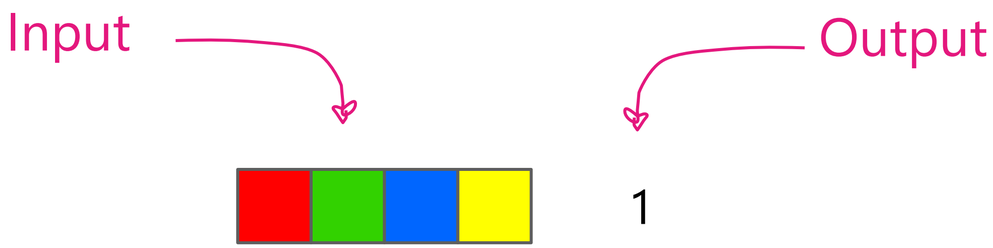
MACHINE LEARNING

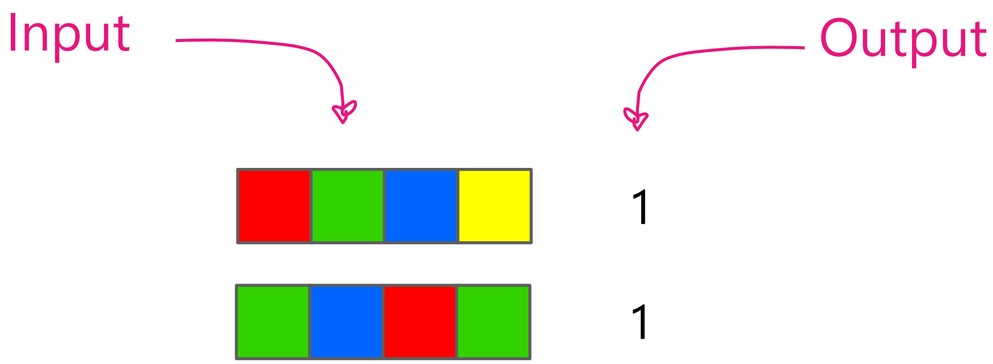


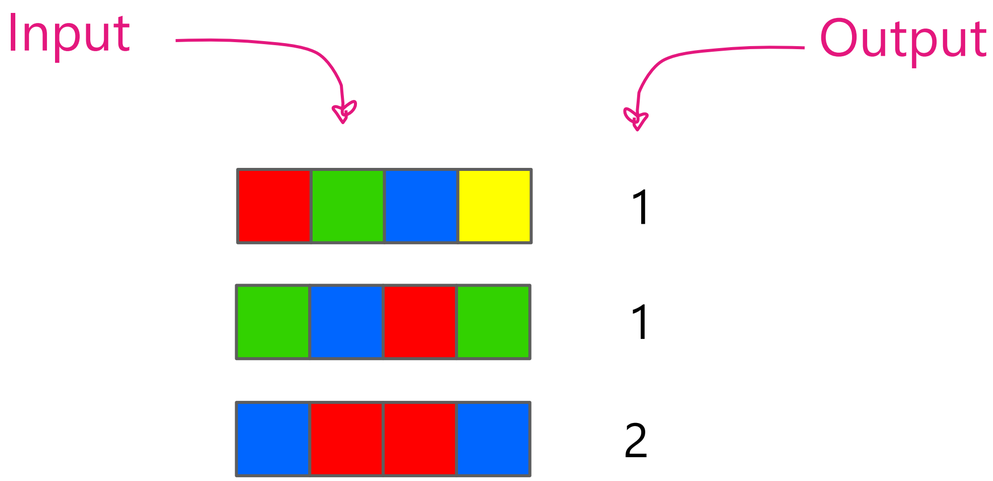
MACHINE LEARNING

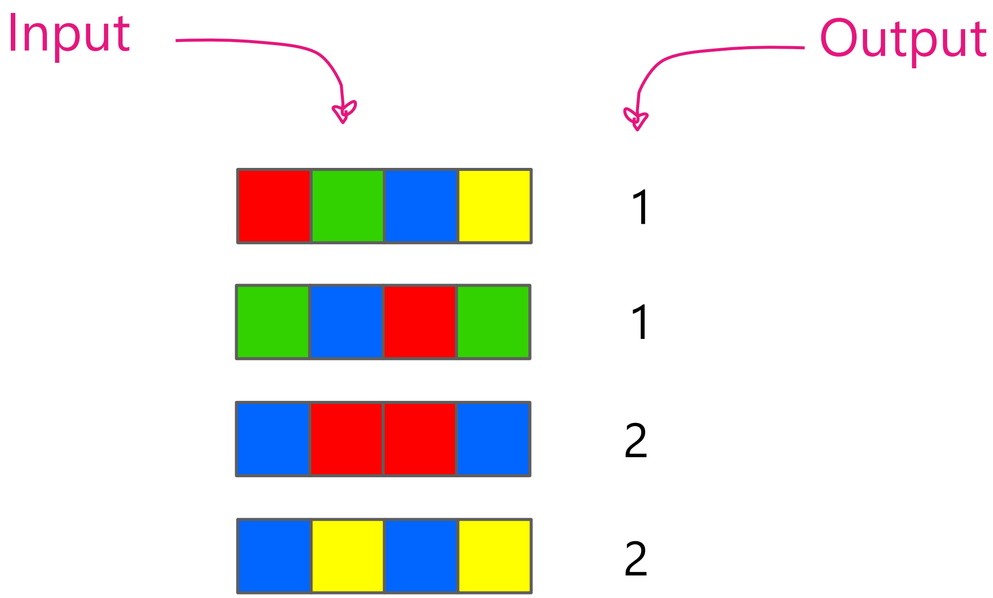


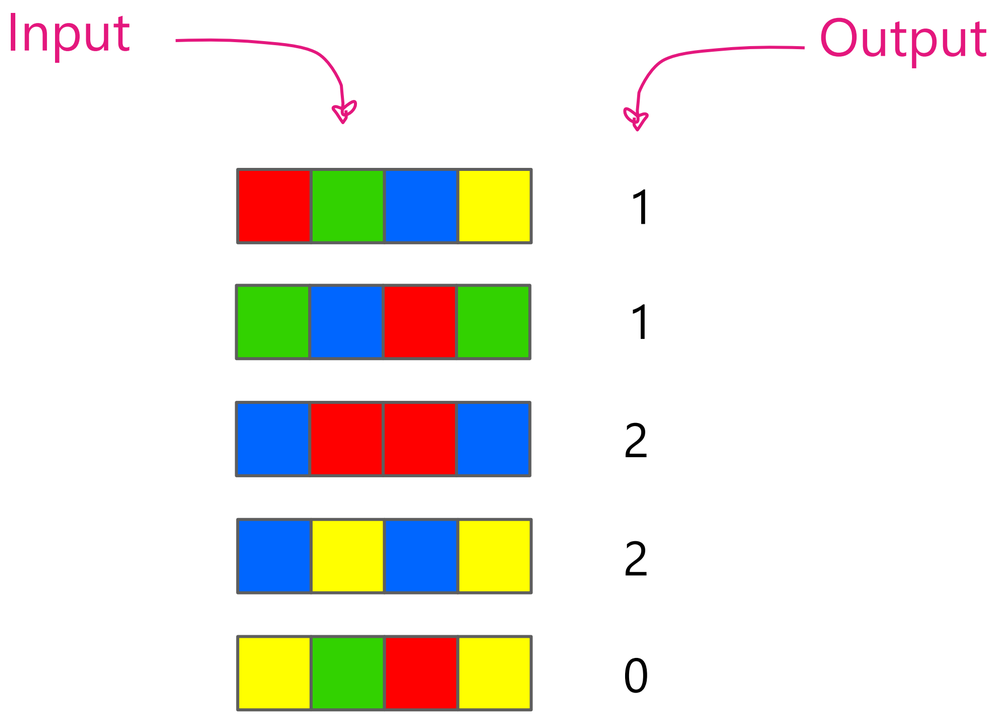
LET’S BUILD A MODEL!

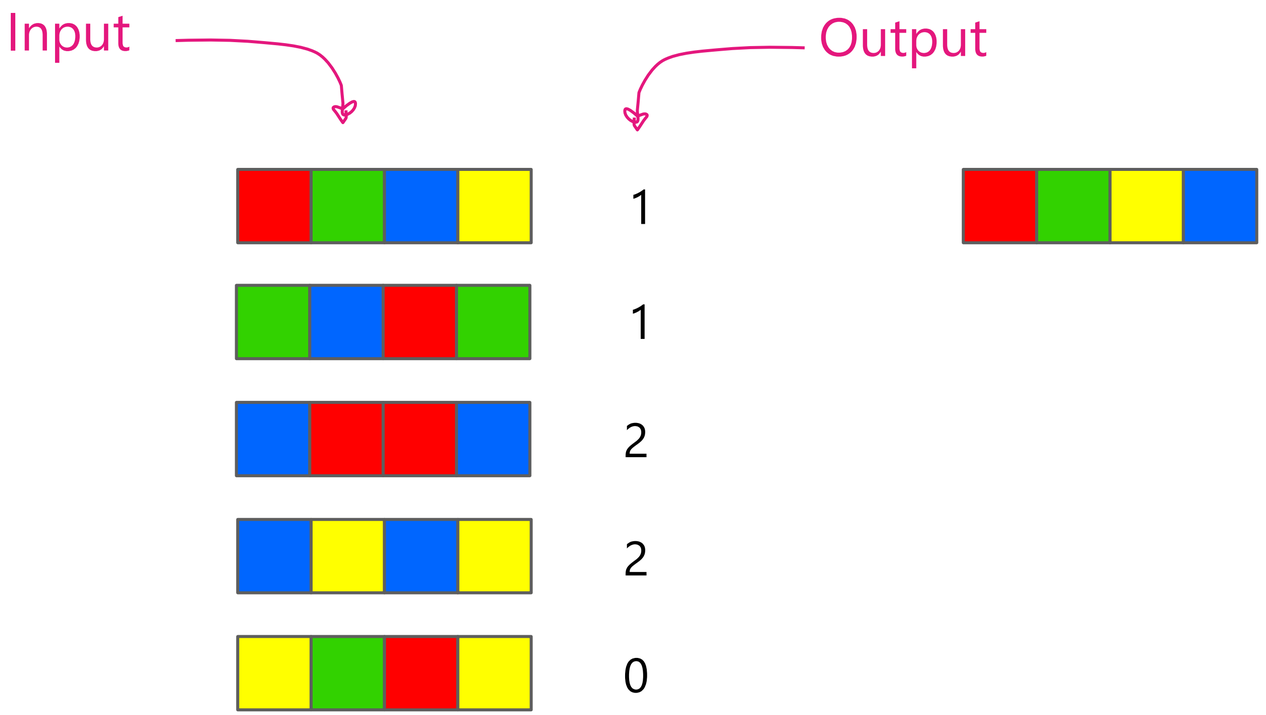

Input
Output
1

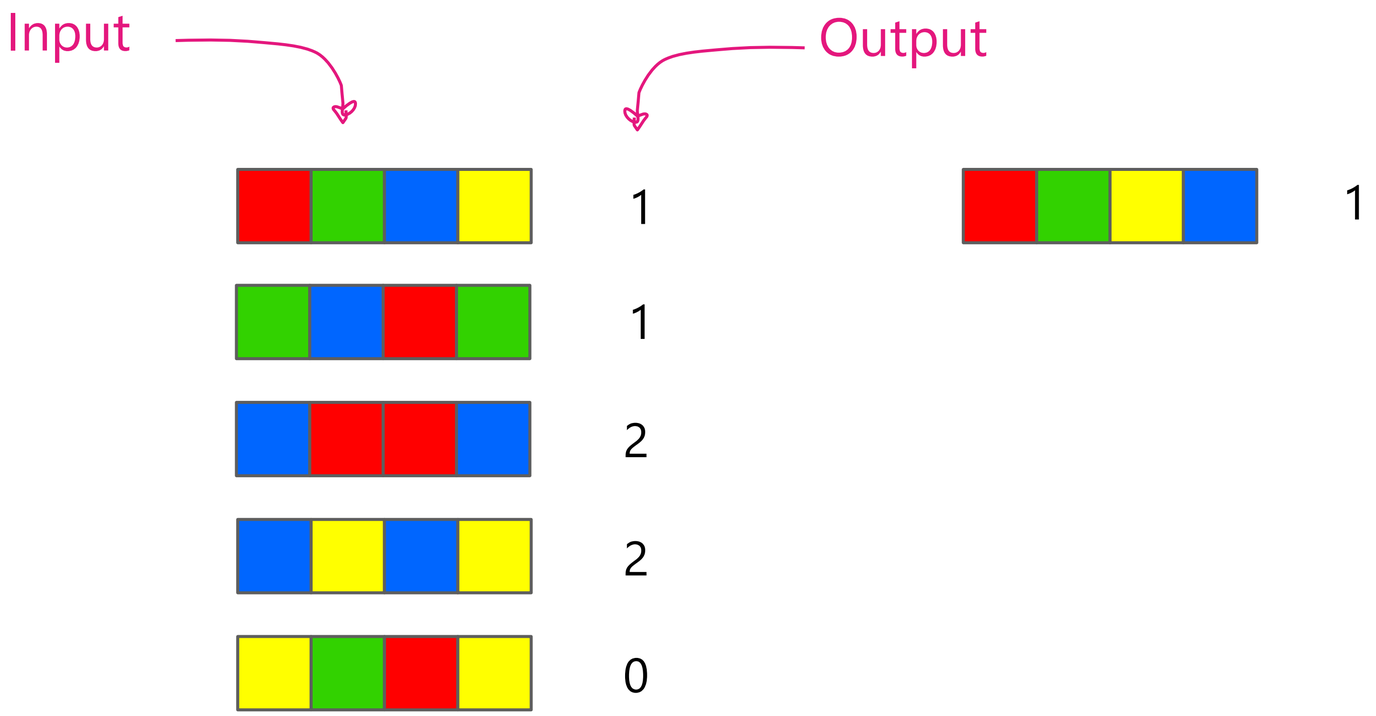

Input
Output
1
1

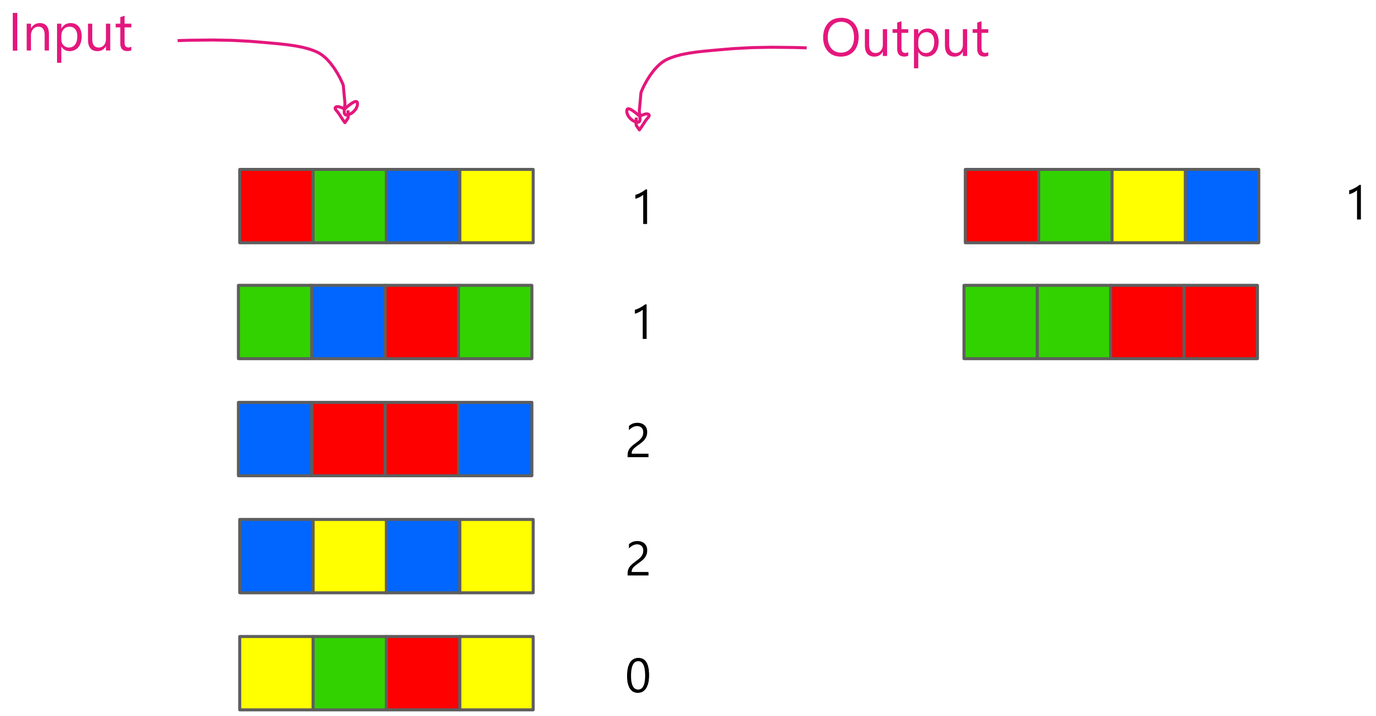

Input
Output
1
1
2

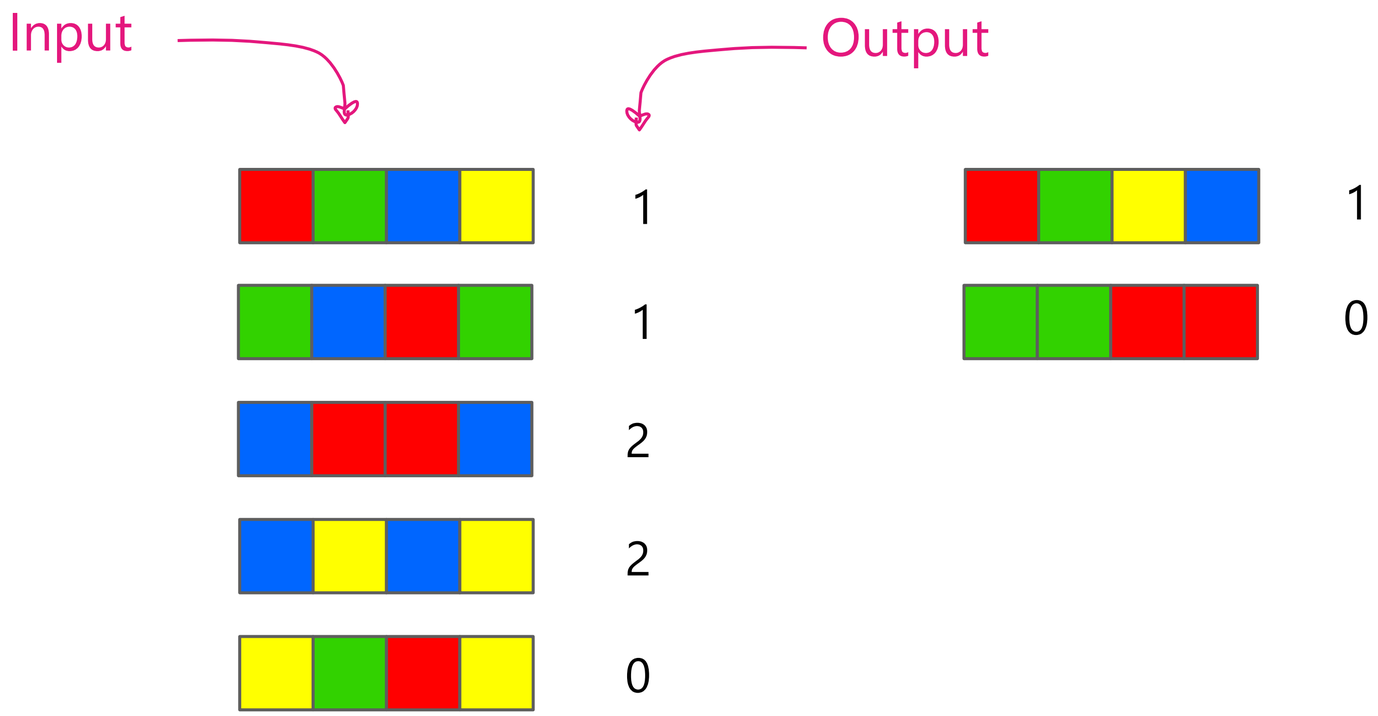

Input
Output
1
1
2
2

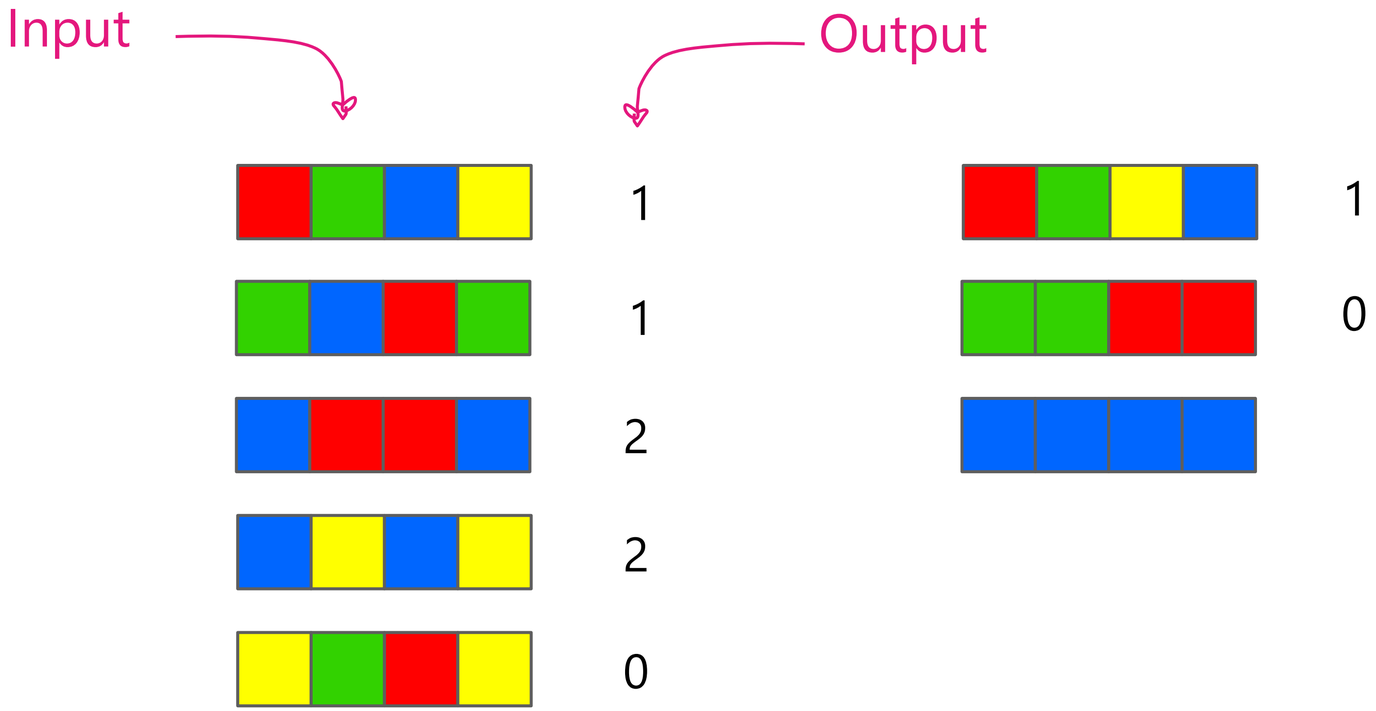

Input
Output
1
1
2
2
0

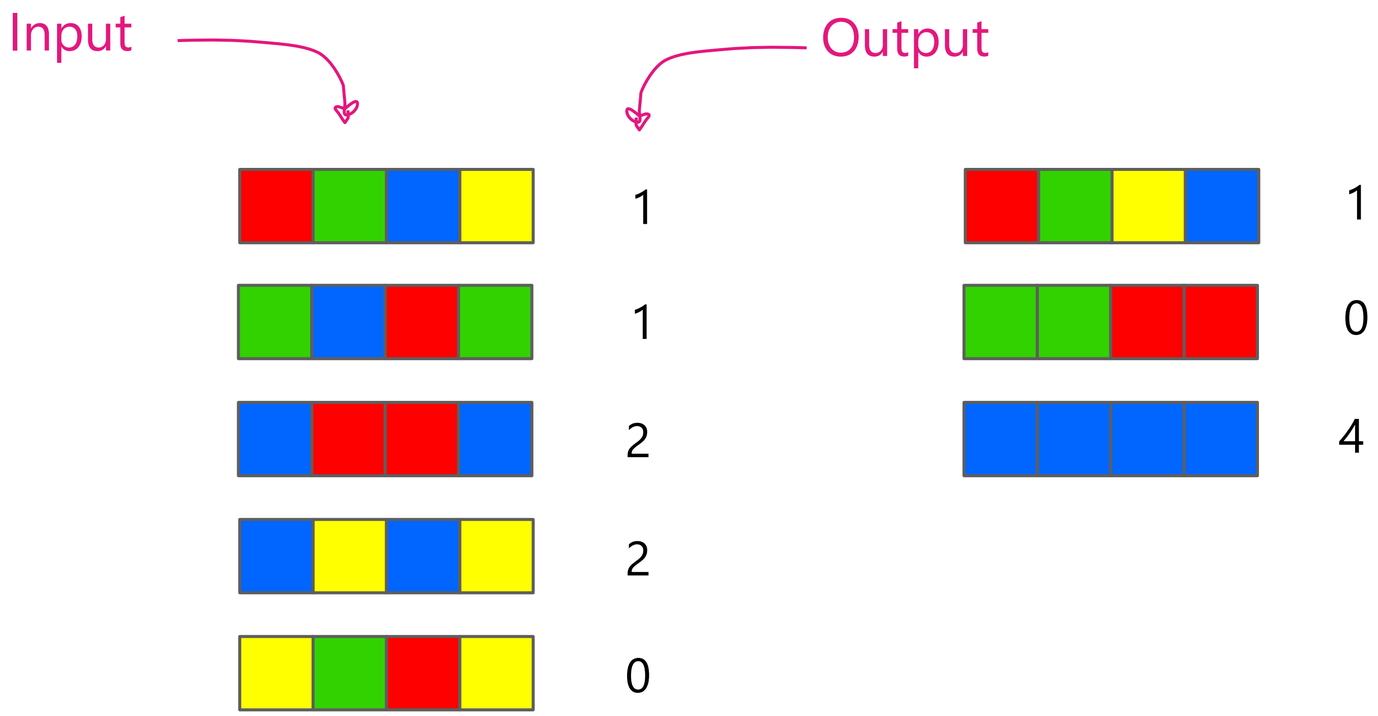

Input
Output
1
1
2
2
0


Input
Output
1
1
2
2
0
1


Input
Output
1
1
2
2
0
1
1


Input
Output
1
1
2
2
0
1
0


Input
Output
1
1
2
2
0
1
0


Input
Output
1
1
2
2
0
1
0
4

LANGUAGE MODELS

LANGUAGE MODELS

A language model is a probability distribution over sequences of words.

Language models can:

1. assign probabilities to sequences of words, and

2. predict the next word and generate sequences.

1) ASSIGN PROBABILITIES

He is *always* late.

He *always* is late.

1) ASSIGN PROBABILITIES

He is *always* late.

He *always* is late.

1) ASSIGN PROBABILITIES

He is *always* late.

He *always* is late.

In his *nineties*, he was still active.

In his *nine teeth*, he was still active.

In his *nine tease*, he was still active.

1) ASSIGN PROBABILITIES

He is *always* late.

He *always* is late.

In his *nineties*, he was still active.

In his *nine teeth*, he was still active.

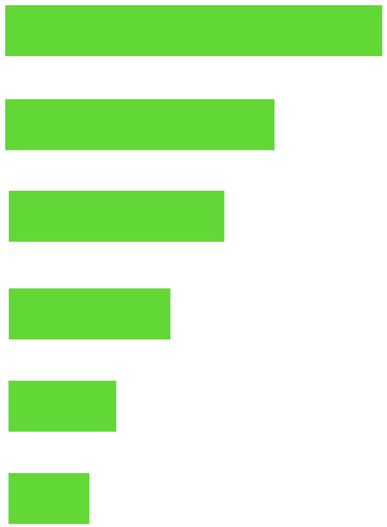
In his *nine tease*, he was still active.

2) PREDICT THE NEXT WORD

I look forward to ...

2) PREDICT THE NEXT WORD

I look forward to



2) PREDICT THE NEXT WORD

I look forward to hearing …

2) PREDICT THE NEXT WORD

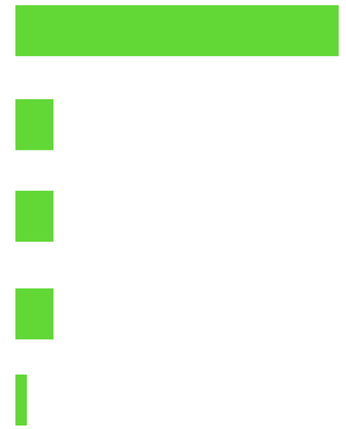
I look forward to hearing

2) PREDICT THE NEXT WORD

I look forward to hearing from ...

2) PREDICT THE NEXT WORD

I look forward to hearing from



LARGE LANGUAGE MODELS (LLMs)

LLMs are language models trained on massive amounts of data using deep neural networks.

They are the basis of what is now known as **Generative AI.**

GENERATIVE AI

GENERATIVE AI

Multimodal Large Language Models (MLLMs) can process and general textual and audiovisual content from a set of instructions (*prompts*). These models are trained on massive amounts of data using deep learning (very big neural networks).

GENERATIVE AI

Using the right *prompts*, we can create practically any kind of content to a high standard.

Writing effective prompts is known as *prompt engineering*. There are a few frameworks to do this (CoSTAR, SMART, CRISP, PAST, etc.).

Example prompt directory: https://prompts.chat/

CREATING EDUCATIONAL MATERIALS

Generative AI is the engine behind many content creation and language assessment tools.

TASK: CONTENT CREATION

Generate a 200-word text passage about the invention of the radio for students at the C1 CEFR level. Next, generate a suitable word list, a true/false exercise (with 5 items), and a multiple-choice cloze test (with 10 items and 4 options each).

What is the quality like? Do they need further editing?

TASK AUTOMATION

We can automate tasks (such as creating content or practising a language) by storing and reusing prompts.

TASK: TASK AUTOMATION

Create a *gem* in Google Gemini to automate the creation of text passages and associated reading comprehension exercises, given a topic, number of words and CEFR level.

ADAPTATION

TASK: TEXT ADAPTATION

Create a new *gem* that takes a newspaper article and generates an adapted version of the text for students at different CEFR levels.

CREATING IMAGES

TASK: CREATING A POSTER

Create a poster to teach A1 learners how to tell the time. The poster must show a big clock and appropriate phrases.

Is it accurate? Can you change details of the image, such as the style of the clock?

CREATING VIDEOS

TASK: CREATING A SHORT VIDEO

Using an existing GPT (e.g. Invideo) on ChatGPT, create a 30-second fun video about homophones in English for B2 learners.

AUTOMATED ESSAY SCORING

AUTOMATED ESSAY SCORING

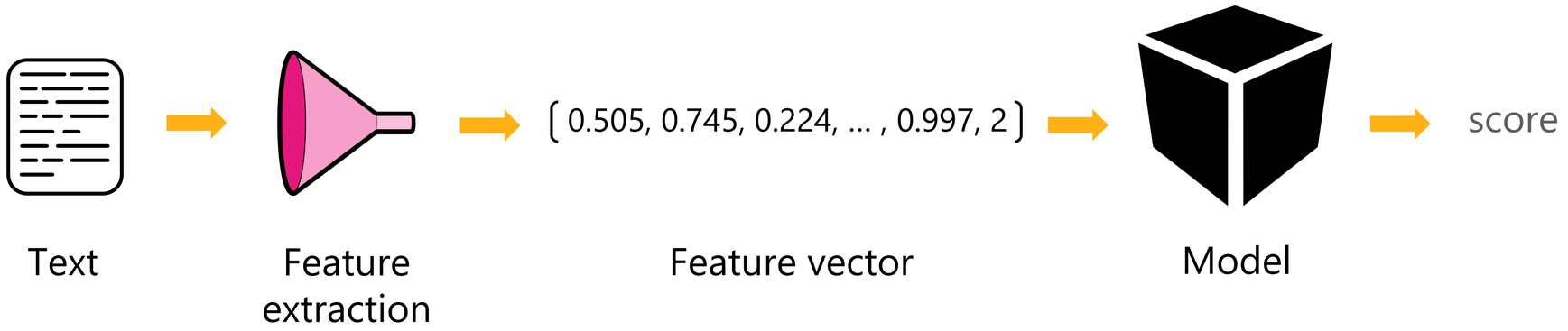
AES models assign a grade to a piece of text.

Grades can be a score, a category (e.g. a CEFR level) or a binary outcome (e.g. pass/fail).

AUTOMATED ESSAY SCORING

AES systems typically score a text on multiple dimensions.

Essays are represented as vectors of numerical features.



SAMPLE FEATURES

Content

- Word overlap with prompt
- Similarity to on-topic essays

Language

- Vocabulary frequency
- Type-token ratio
- Error rates
- Average language model sentence probability

I don’t think that teenagers are too young to teach other people about anything. Of course, they can’t know very well some things, for example: some scientific theories, history, mathematic at all and etc, but a lot of teenagers know a lot about technology. It’s normal for them to spend a lot of time with computer, different gadgets. Most of them know, how these gadgets work, so they can explain other people different moments of their working. My Granny often asks to me for a piece of advice about her mobile phone.
Teenagers’ knowledge about technology usually based on practice, so often they don’t know about process of creating the phone, the TV, etc. They really shouldn’t try to tell about things, which they don’t know.

Structure

- Essay length
- Words per sentence
- Clauses per sentence
- Average sentence length
- Number of paragraphs
- Cohesion
- Coherence

AUTOMATED ESSAY SCORING

Today, Generative AI is also used to assess essays and provide feedback. However, it is often less transparent than feature-based systems – it's a *black box*.

TASK: FEATURE-BASED AES VS. LLMS

Assign a CEFR level to a sample student essay using a feature-based AES system (e.g. Write&Improve) and an LLM. Which of them seems more accurate?

AUTOMATED ESSAY SCORING

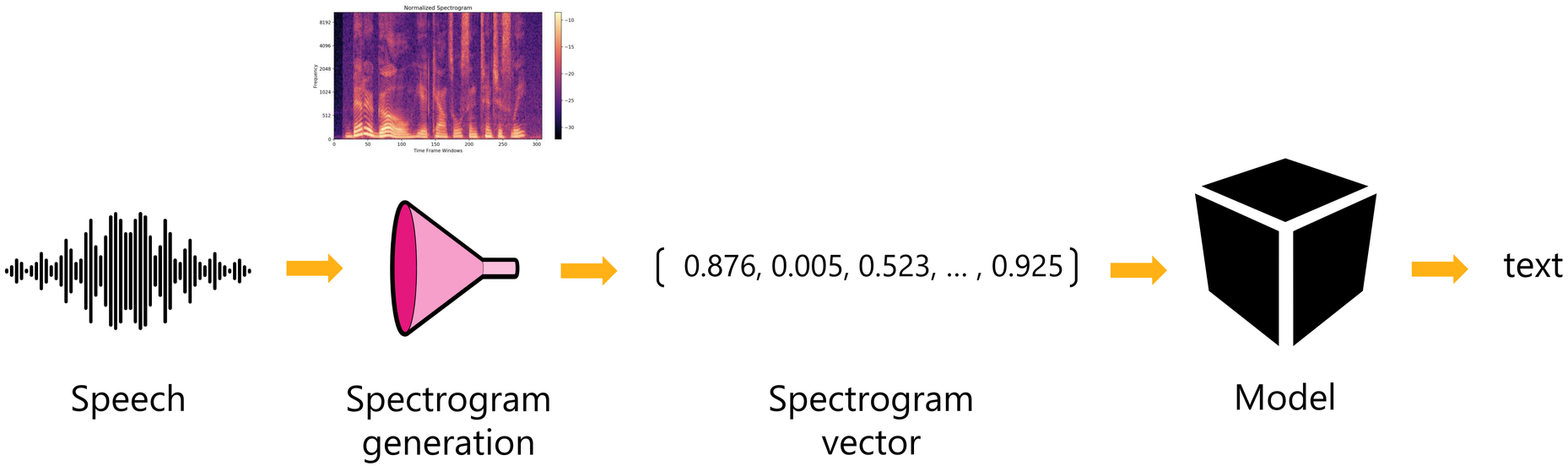
DISCUSSION

AES can be controversial. What are the risks? In what scenarios would it be/not be appropriate?

AUTOMATIC SPEECH RECOGNITION

AUTOMATIC SPEECH RECOGNITION

ASR systems “understand” human speech and convert it to text.



AUTOMATIC SPEECH RECOGNITION

The accuracy of ASR systems is measured using Word Error Rate. WER is defined as the number of errors in the transcription divided by the number of words in the reference.

ASR I like to choke a lot
REF I like to joke a lot

WER = 1/6 = 16.66%

AUTOMATIC SPEECH RECOGNITION

TASK: EVALUATING ASR SYSTEMS

1. Listen to the following audio clip and manually transcribe what you hear.

2. Use an ASR service to generate an automatic transcription of the audio.

3. Compare both transcriptions and compute WER. Which ASR system is better?

AUTOMATIC SPEECH RECOGNITION

TASK: AUTOMATIC SPEAKING ASSESSMENT

1. Using an LLM, write a prompt to perform automatic speaking performance assessment on three aspects: fluency, language use and pronunciation. Ask the LLM to try to guess the learner's L1 and provide reasons for its decision.

2. Upload an audio file or (record yourself reading aloud). Is the assessment accurate?

AUTOMATIC SPEECH RECOGNITION

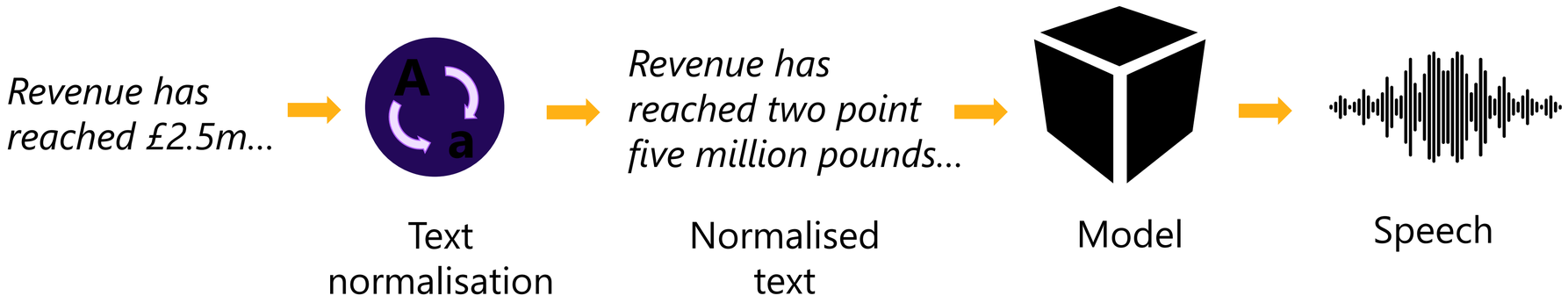
DISCUSSION

How could we use ASR technology for assessment?
What problems could we encounter?

SPEECH SYNTHESIS
(TEXT TO SPEECH)

SPEECH SYNTHESIS (TEXT TO SPEECH)

Text-to-speech systems convert written text into speech.
Modern systems use very realistic voices.



SPEECH SYNTHESIS (TEXT TO SPEECH)

TASK: IMPROVING PRONUNCIATION

Using an LLM, create sentences that focus on common aspects of pronunciation that learners often find difficult (e.g. short vs long vowels). Next, generate model spoken versions using a TTS service.

SPEECH SYNTHESIS (TEXT TO SPEECH)

TASK: CREATING YOUR OWN SPEAKING SCRIPTS

Create a listening comprehension question consisting of a dialogue between three people with different accents.

1. Create a short dialogue and a question.

2. Create an audio file for each turn of the conversation using a TTS service.

3. Join all the audio files together.

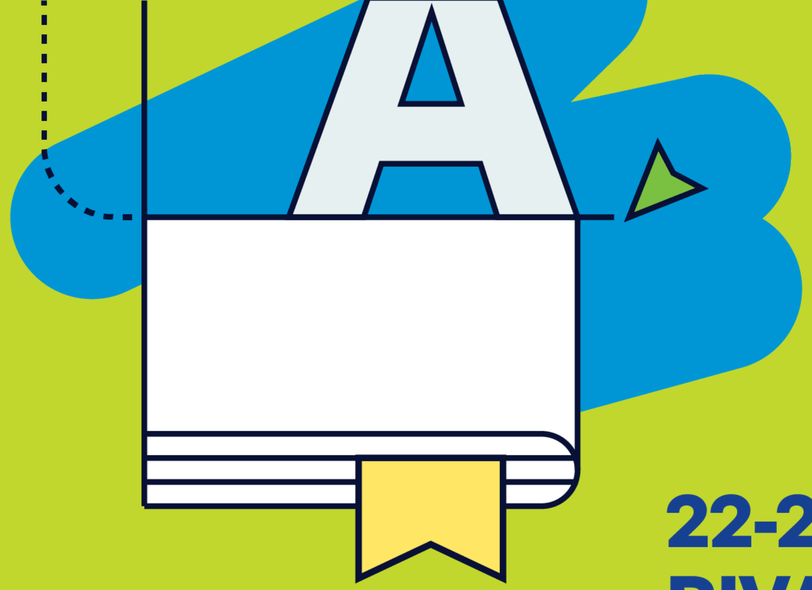
FINAL WORDS

- Learn, evaluate and use.
- Use AI responsibly and ethically.
- Be creative – the sky's the limit!
- Explore! There are lots of tools: https://theresanaiforthat.com/
- Plenty of resources for teachers from the British Council (see links online).

THAT'S ALL FOLKS!

mariano.felice@britishcouncil.org

didacta italia EDIZIONE TRENTINO



22-24 OTTOBRE 2025
RIVA DEL GARDA

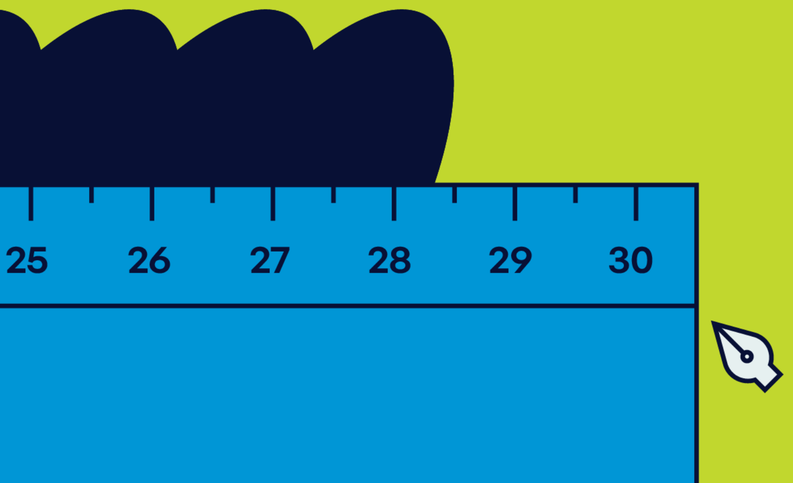
Finanziato da

COESIONE ITALIA 21-27 TRENTINO

Cofinanziato dall'Unione europea

PROVINCIA AUTONOMA DI TRENTO

Iniziativa realizzata nell'ambito del Programma FSE+ 2021-2027 della Provincia autonoma di Trento, con il cofinanziamento dell'Unione europea -Fondo sociale europeo plus, dello Stato italiano e della Provincia autonoma di Trento



Organizzazione

FIRENZE FIERA

Partner Scientifico

INDIRE ISTITUTO NAZIONALE DOCUMENTAZIONE INNOVAZIONE RICERCA EDUCATIVA

IPRASE

Partner

didacta international

Comitato organizzatore

PROVINCIA AUTONOMA DI TRENTO

Autonome Provinz Bozen Provincia autonoma di Bolzano Provincia autonoma de Bulsan SÜDTIROL · ALTO ADIGE

USR FVG

USRLIGURIA UFFICIO SCOLASTICO REGIONALE

Ufficio Scolastico per la Lombardia

Regione Autonome Vallée d'Aoste Regione Autonoma Valle d'Aosta

MIM UFFICIO SCOLASTICO REGIONALE PER IL VENETO Direzione Generale

Sponsor Gold

wacebo

Sponsor Silver

BRITISH COUNCIL

C2 GROUP

CampuStore

MR*DIGITAL EDUCATION

Supporter

C&C