**Week1**

A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

这里A说的不是Dimension，而是向量的长度

cosine similarity returns 1 (maximum similarity) for vectors pointing in the same direction, regardless of their magnitudes. Thus, it fails to distinguish such cases.

A white background with black text

AI-generated content may be incorrect.

A black background with white text

AI-generated content may be incorrect.

A black background with white text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer

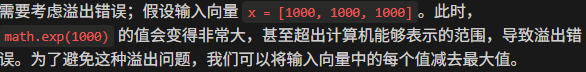
AI-generated content may be incorrect.

A black background with white text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

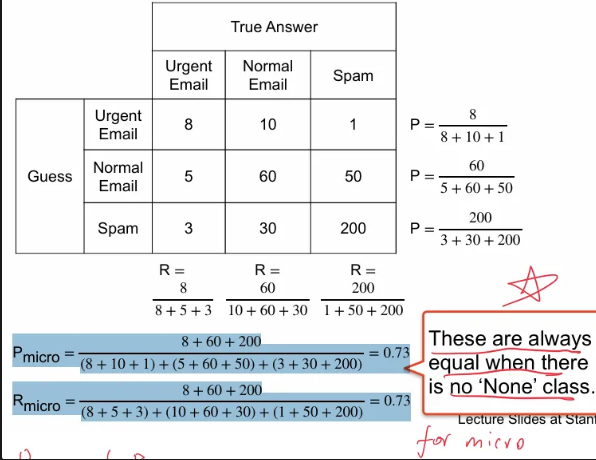


**Week2**

A screenshot of a computer

AI-generated content may be incorrect.

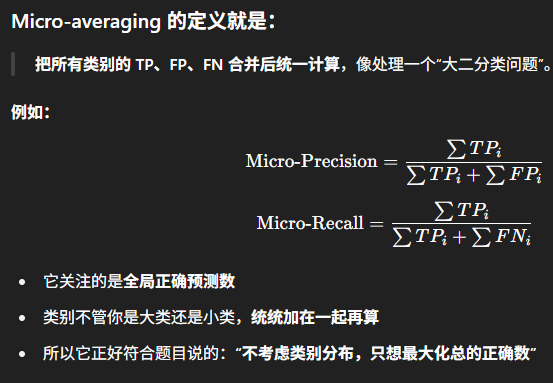
Micro公式中你会发现把所有class当整体来计算



A screenshot of a computer screen

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect. 

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A black screen with white text

AI-generated content may be incorrect.

即只关注正类的表现，不关注负类的表现

A white background with black text

AI-generated content may be incorrect.

**Week3**

**A white background with black text

AI-generated content may be incorrect.**

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect. A black background with white text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A black background with white text

AI-generated content may be incorrect.

A screenshot of a cell phone

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.



A screen shot of a black background

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A close up of a text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

**用 micro 就能更好衡量整个模型对所有 token 的正确性，而不是被小类扰动**

**Week4**

A screenshot of a computer

AI-generated content may be incorrect.

在候选词中，既考虑语言模型（LM）的概率，又考虑候选词插入到当前上下文后，整体句子的向量表示与上下文的一致性（similarity）

A white background with black text

AI-generated content may be incorrect.

这一题只有TOP-P是要用累加概率达到threshold

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

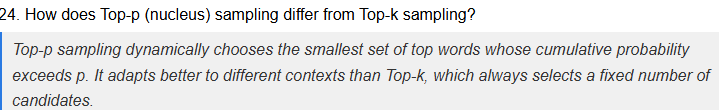
A black background with white text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

图结构解码适用于结构化预测，是因为它能明确建模部分输出之间的依赖关系。



**Week5**

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.





A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

 **A. They cannot generate outputs**：LLMs肯定能生成输出，不正确。

 **C. They are too fast**：速度不是导致不可靠的原因。

 **D. They use rule-based logic**：LLMs不是基于规则逻辑，而是基于统计学习。

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.





A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

比如可能有很多个人叫做张三

A screenshot of a white screen

AI-generated content may be incorrect.

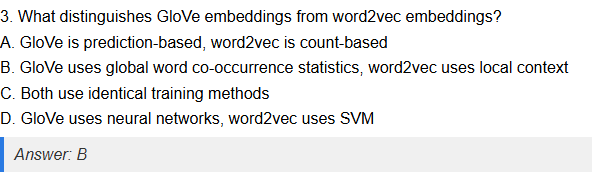
A screenshot of a white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

**Week6**



A black background with white text

AI-generated content may be incorrect.

A screenshot of a computer

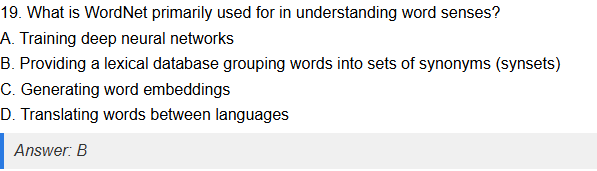
AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.



WordNet 是一个**基于词义组织词汇的语义网络**，主要用来理解词语的意思和它们之间的关系，不是一个训练工具或翻译器。Synonyms是同义词的意思

A close-up of a computer screen

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A close up of a white background

AI-generated content may be incorrect.



BPE 会“机械性地合”，最后可能得到一些**没有语义独立性的组合**，因为它**不优化全局分词质量**，只是贪心合并。而Unigram会保留常见的词

**Week7**

A white background with black text

AI-generated content may be incorrect.

A screenshot of a black background with white text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.



Multiplicative Attention不包含非线性

Additive Attention包含非线性变化

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

 A（它不是通过训练学习的）是事实，也是sinusoidal positional encoding的缺点之一，因为不能根据数据优化位置编码。

 B（它不能泛化到比训练时更长的序列）是不正确的说法，sinusoidal编码的设计初衷之一就是可以推广到比训练更长的序列，因为它是基于连续的正弦函数，可以计算任意位置的编码。（Learnable Positional Embedding不行）

 C（它用固定的正弦函数编码位置）也是事实，也是缺点，因为固定形式可能限制表达力。

 D（可能无法捕捉复杂的位置模式）也是缺点，因为它是预定义函数，可能不足以捕获复杂的位置信息。

A screenshot of a computer

AI-generated content may be incorrect.

Self-attention 计算是基于矩阵乘法和softmax函数，梯度可以通过标准的链式法则直接计算，因此用**直接对矩阵乘法和softmax求导**。

A white background with black text

AI-generated content may be incorrect.

A close up of a text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

Smooth gradient

A screenshot of a white background

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

A close-up of a text

AI-generated content may be incorrect.

这里主要看看layer normalization，有点不熟悉



因为 decoder 需要根据 encoder 的输出（即编码过的输入序列）来生成输出。

A white background with black text

AI-generated content may be incorrect.

这里就是self-attention对比hidden layer

**Week8**

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

And also WER



how *confused* or *uncertain* a language model is when predicting a sequence of tokens.

* **Lower perplexity is better** — it means the model is more confident in its predictions.
* **Minimum value is 1** — when the model predicts every token with probability 1.



how well a system ranks the correct answer **across multiple queries**, the closer to 1 the better.



**Intrinsic metrics** assess the model’s language ability (Perplexity, MRR) directly, while **extrinsic metrics** (BLEU, WER) evaluate its impact on real-world tasks.

**Week9**

A white background with black text

AI-generated content may be incorrect.

A close up of a text

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

The key idea here is that in-context learning need to find pattern, and a well organized structure can help them find patterns, but labels is not that important when we do the enough training.

**Week10**

A screenshot of a computer

AI-generated content may be incorrect.

**Trafilatura** 是一个专门用于**网页文本提取**的 Python 工具库，它的主要用途是：

**从网页中提取高质量的正文内容**，去除广告、导航栏、脚注等无关内容。

A screenshot of a computer

AI-generated content may be incorrect.

**Week11**

A screenshot of a question

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

**对于 instruction tuning 来说，答案是：**  
**No, increasing training examples does *not* always improve generation quality.**

对于 **In-Context Learning**，答案通常是：**Yes** —— 增加示例数通常可以**提升生成质量**，尤其在 few-shot 设置下。

A white background with black text

AI-generated content may be incorrect.



Different sentence may have same score in token-level, but their meaning is different. That’s why we need perplexity to help model really understand question.



they don't always reflect true human preferences.

A black text on a white background

AI-generated content may be incorrect.

A white background with black text

AI-generated content may be incorrect.

**Week12**

A white background with black text

AI-generated content may be incorrect.

A close up of a white background

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.