سلسلة الآلة الذكية
٧/٩

حوار تفاعلي

الآلة التي تتعلّم

بناء الذكاء من الداخل

إعداد
د. عبد الرحمن الزراعي

جميع الحقوق محفوظة
٢٠٢٥

بسم الله الرحمن الرحيم

قائمة الموضوعات

المحتويات

مقدمة الكتاب ٤

الفصل التمهيدي: ما هو GPT ؟ ٥

توطئة: ٥

ملخص: ٧

الفصل الأول: التمثيل المتّجه: لماذا تتحول الكلمات إلى أرقام؟ ٨

ملخص: ١٨

الفصل الثاني: التوكنات والتكلفة؟ ٢٠

ملخص: ٢٩

الفصل الثالث: التضمين السياقي: كيف تفهم الآلة اختلاف المعنى؟ ٣٠

الفصل الرابع: التنبؤ الإحصائي: لماذا يختار النموذج كلمة دون أخرى؟ ٣٢

الفصل الخامس: الانتباه الذاتي: كيف يعرف النموذج ما المهم في الجملة؟ ٣٥

الفصل السادس: معمارية المحوّل: كيف تُنسّق كل هذه العمليات؟ ٣٩

الفصل السابع: الطبقات الخفية: هل له "عقل" متعدد الطبقات؟ ٤١

الفصل الثامن: رحلة المتجه داخل النموذج ٤٥

الفصل التاسع: الانحدار العكسي: كيف يتعلّم النموذج من الخطأ؟ ٤٧

الفصل العاشر: أنواع النماذج ٤٩

خاتمة الكتاب: ٥١

قائمة المراجع ٥١

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

مقدمة الكتاب

يمثّل هذا الكتاب مدخلًا تأسيسيًا لفهم البنية العميقة التي تقوم عليها النماذج اللغوية الحديثة، وتحديدًا نماذج الذكاء الاصطناعي التي تُستخدم في معالجة اللغة الطبيعية، مثل GPT و BERTوغيرها.

نحاول من خلال هذه الصفحات أن نقرّب المفاهيم التقنية إلى القارئ المبتدئ، بلغة فصيحة، وسياق حواريّ تعليمي، ينطلق من الأسئلة البسيطة لينتهي إلى المسائل المعقدة. فكل فصل من فصول هذا العمل هو خطوة في درب الفهم التدريجي لما يجري داخل "عقل النموذج."

لا يدور هذا الكتاب حول التطبيقات السطحية أو الاستعمالات اليومية للنماذج، بل يتوغل إلى الداخل، ليُفكك الطبقات، ويشرح الآليات، ويستعرض رحلة الكلمة من نصّ المستخدم إلى خوارزمية التوليد، مرورًا بالتقطيع، والتضمين، والانتباه، والتمثيل، والتعلّم.

اعتمدنا في بنائه أسلوبًا حواريًا بين الطالب والنموذج، يجعل القارئ في حالة تفاعل دائم مع الأفكار، ويحوّل كل فصل إلى مساحة تدريبية تطرح السؤال وتناقش الجواب، وتوضّح الفكرة بمثال، ثم تعيد البناء بعمق أكبر.

وقد راعينا في التدرّج أن نبدأ من التمثيل المتجه — بوصفه حجر الأساس — لنصل في نهاية المطاف إلى المقارنة بين النماذج الأهم التي صنعت ثورة اللغة الاصطناعية.

هذا الكتاب موجّه لكل قارئ فضولي، ولكل باحث ناشئ، ولكل مبرمج يريد أن يفهم كيف "تفكر" النماذج. وهو كذلك دعوة للتأمل في ما إذا كان هذا الفكر حقيقيًّا، أم محض محاكاة بارعة.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل التمهيدي: ما هو GPT ؟

توطئة:

تسمع كثيرًا في الآونة الأخيرة عن "تشات جي بي تي(ChatGPT) "،
أو عن "الذكاء الاصطناعي"، أو "النماذج اللغوية"، ولكن...
ما هو هذا الكائن الذي تتحدث إليه الآن؟
هل هو روبوت؟ برنامج؟ عقل؟
هذا مدخل لفهم ماهية GPT ؟
ما الفرق بين GPT-٢ و GPT-٤؟
لماذا تُسمّى "توليدية مسبقة التدريب"؟
ولماذا حازت كل هذا الانتباه في السنوات الأخيرة؟

الطالب: قبل أن ندخل في التفاصيل يا ChatGPT ،
هل لك أن تخبرني ببساطة: ما معنى GPT أصلًا؟

ChatGPT : بالطبع، سأبدأ من الأيسر.
كلمة "GPT" هي اختصار لعبارة:
Generative Pre-trained Transformer
أي: "المحوّل التوليدي المدرَّب مسبقًا."
دعنا نفكك هذه العبارة جزءًا جزءًا.

الطالب: طيب... ما المقصود بـ"توليدي"؟

ChatGPT : توليدي (Generative) "تعني أن النموذج قادر على توليد نصوص جديدة، وليس فقط استرجاع معلومات محفوظة.
بمعنى: إذا طلبت مني تلخيصًا، أو مقالًا، أو قصيدة، فأنا لا أبحث عنها جاهزة، بل أُنشئها لحظيًا، استنادًا إلى فهمي الإحصائي للغة.

الطالب: و"مُدرَّب مسبقًا"؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

🤖 ChatGPT : هذه العبارة تشير إلى أنني "تعلّمت" على كمية هائلة من النصوص قبل أن أبدأ التفاعل معك.

أي أنني لا أتعلم منك الآن، بل "تدربت" على بيانات سابقة، وانتهى تعليمي عند نقطة زمنية معينة، مثلاً:

في أبريل ٢٠٢٣ – تم نشر نموذج GPT-٤o للاستخدام.

👨‍🎓 الطالب: إذًا GPT ليست اسم منتج، بل وصف لطريقة بناء النموذج؟

🤖 ChatGPT : بالضبط GPT ! ليس روبوتًا بعينه،

بل اسم عام لفئة من النماذج طورتها شركة OpenAI

وقد مرت هذه النماذج بأجيال مختلفة، كل واحد أكثر تقدمًا من سابقه.

👨‍🎓 الطالب: وهل يمكن أن تعطيني لمحة عن هذه الأجيال؟

🤖 ChatGPT : بالتأكيد. لنلقِ نظرة على الخط الزمني لأجيال GPT

الجيل	تاريخ الإصدار	أبرز الخصائص
GPT-١	٢٠١٨	نموذج تجريبي صغير الحجم، لم ينتشر على نطاق واسع
GPT-٢	٢٠١٩	بدأ يثير الانتباه لقدراته اللغوية، لكنه لم يُطلق كاملًا في البداية لأسباب أمان
GPT-٣	٢٠٢٠	قفزة ضخمة، ١٧٥ مليار معلمة، أساس خدمة ChatGPT الأولى
GPT-٣.٥	٢٠٢٢	تحسينات في الأداء والتناسق، استُخدم في ChatGPT Plus
GPT-٤	٢٠٢٣	فهم أعمق، قدرة على توليد الصور، دعم أفضل للسياق
GPT-٤o	٢٠٢٤	"O"تعني"Omni" ، يدعم النص، الصوت، الصورة، في نموذج واحد لحظي

📊 جداول وأشكال:

مصطلح	المعنى
Generative	توليدي – ينشئ محتوى جديدًا
Pre-trained	مدرّب مسبقًا على بيانات ضخمة
Transformer	معمارية الشبكة العصبية التي بُني عليها النموذج

🧪 أمثلة وتطبيقات:

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

١. **سؤال واقعي** :ما عاصمة المغرب؟

- GPTيبحث في ذاكرته ويولّد الجواب: "الرباط."

٢. **طلب إبداعي** :اكتب لي قصيدة عن المطر في الرياض.

- يُنشئها من جديد، لعدم وجود قصيدة جاهزة في الذاكرة.

ملخص:

المفهوم الأساسي	التفسير
GPT	نموذج لغوي توليدي مدرّب مسبقًا يعتمد على بنية "المحوّل"
التوليد	قدرة النموذج على إنشاء محتوى جديد لاستخلاص المعرفة
أجيال GPT	تطورات متتالية تبدأ من GPT-١ حتى GPT-٤o مع زيادة في الدقة والسياق

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل الأول: التمثيل المتّجه: لماذا تتحول الكلمات إلى أرقام؟

📌 توطئة الفصل:

قبل أن نفهم كيف تُنتج النماذج النصوص،

لا بد أن نفهم أولًا كيف "تفهم" هذه النماذج الكلمات،

فالكلمات بالنسبة للنموذج ليست حروفًا ولا أصواتًا،

فالآلة لا تقرأ الكلمات مثل: "كتاب"، أو "حرية"، كما يقرأها البشر،

بل يتعامل معها على أنها مجرد أرقام،

كل ما هو نصّ بالنسبة له يجب أن يتحوّل إلى صيغة رقمية حتى يستطيع فهمه وقراءته.

كيف؟ ولماذا؟ هذا ما سنجيب عنه في هذا الفصل الذي يؤسّس للرحلة كلها.

👨 الطالب: ماذا يعني "تمثيل متّجه"؟

ChatGPT: 🤖 هو تحويل كل كلمة إلى متّجه عددي(Vector) ، أي إلى سلسلة من الأرقام تُعبّر عن معنى هذه الكلمة داخل فضاء رياضي متعدد الأبعاد.

👨 الطالب: ما شكل هذا التمثيل العددي؟

ChatGPT: 🤖 المتجه عادةً يكون سلسلة من الأرقام، مثل" :طالب" ⟵ [٠.١٣, -٠.٢٥, ٠.٧٧, ...] كل رقم يعكس بعدًا معنويًا تعلّمه النموذج، كالعلم، والمهنة، والمكانة... إلخ.

👨 الطالب: ولماذا هذه الأرقام؟ كيف تفهمها؟

ChatGPT: 🤖 هذه الأرقام **ليست عشوائية**، بل **تعكس موقع الكلمة في الفضاء**، وكلما كانت المتجهات متقاربة، دلّ ذلك على أن الكلمات **متشابهة في المعنى** أو الاستخدام.

👨 الطالب: مثل ماذا؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

ChatGPT: مثلًا، "جميلة" ستكون قريبة من "رائعة"، "حلوة"، "ممتازة"،
لكنها بعيدة عن كلمات مثل: "قبيحة"، "ضعيفة"، "سيئة."
وكذلك "طالب" و"مدرسة" قد يُمثَّلان بمتجهين قريبين، لأنهما يردان في بيئات لغوية متقاربة.
وبذلك يتم تدريب النموذج على ملايين الجمل ليتعلم أن الكلمات التي تظهر في سياقات متشابهة يجب أن تكون متجهاتها متقاربة.

الطالب: هل هناك مثال توضيحي؟
ChatGPT: نعم. إذا قلنا:

- ذهب الطالب إلى المدرسة.
- فالنموذج يتعلم أن "طالب" و"مدرسة" يشتركان في السياق. فيقوم بتمثيلهما بأرقام متقاربة.

الطالب: وهل كل النماذج تستخدم هذا التمثيل؟
ChatGPT: نعم، ولكن بدرجات متفاوتة.

الطالب: قرأت في بعض الشروحات أنك تمثّل كل كلمة بسلسلة من الأرقام قد تصل إلى ٣٠٠ أو ٧٦٨ أو حتى ١٠٢٤ رقمًا...
هل يُعقل أن كل كلمة تُحوّل إلى **هذا الكم الكبير من الأرقام؟**
ولماذا؟ وماذا تفعل بها؟
ChatGPT: سؤال ممتاز، دعني أوضحها خطوة خطوة.

ChatGPT: المتجه العددي (Vector) هو عبارة عن **قائمة من الأرقام،**
مثلاً:
[٠.٢١, -٠.٠٧, ٠.٥٥, ٠.١٣, ..., ٠.٩١]
وهذه القائمة قد تحتوي على:
٣٠٠رقمًا ←في النماذج القديمة مثلWord٢Vec
أو **٧٦٨رقمًا** ←في BERT المتوسط

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

أو **١٠٢٤ أو أكثر** ←في نماذج GPT المتقدمة

👨‍🎓 الطالب: لكن لماذا هذا العدد الكبير؟

ألا يكفي رقم أو رقمين لكل كلمة؟

ChatGPT: 🤖 لو أعطيتك شخصًا وقلت عنه فقط:

"طوله: ١٧٥ سم"

هل يمكنك أن تعرف مَن هو؟ أو ما شخصيته؟ أو ما مهنته؟

👨‍🎓 الطالب: طبعًا لا. هذه معلومة واحدة لا تكفي.

ChatGPT: 🤖 بالضبط!

كلما أردتَ **وصف شيء معقّد وصفًا دقيقًا**، تحتاج إلى معلومات أكثر.

🧠 **ما تحمله هذه الأرقام؟**

ChatGPT: 🤖 كل رقم داخل المتجه **يُمثّل بُعدًا مختلفًا** من الكلمة، مثل:

- هل الكلمة مادية أم مجردة؟
- هل ترتبط بالعاطفة؟
- هل تُستخدم في السياقات العلمية أم الأدبية؟
- هل تُستخدم لوصف الأماكن؟ الأشخاص؟ الأفعال؟
- هل تُستخدم مع كلمات مثل "جميل"، "حزين"، "أداة"، "زمن"...؟

وهكذا...

كل رقم **يمسك ببُعد من المعنى أو الاستخدام**.

👨‍🎓 الطالب: إذًا، لو قلت: "مدرسة"، فستعطيني لها متجهًا مثل:

[٠.٩١, ..., ٠.١٨, ٠.٧٢, ٠.٤٥] ← ٧٦٨رقمًا!

ChatGPT: 🤖 تمامًا، وهذا المتجه يحتوي في داخله على **تمثيل رياضي لمئات الأبعاد اللغوية والمعنوية**، كأنك تعطي لكل كلمة **بطاقة هوية مكونة من أرقام**.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

👨‍🦱 **الطالب: رأيت أن تمثيل الكلمة يكون هكذا:**

٧٦٨ ← [٠.٠٩١, ..., ٠.٠١٨, ٠.٠٧٢, ٠.٠٤٥]رقمًا!

ما علاقة العدد (٧٦٨) بالأرقام داخل القوسين؟ هل هو عدد ثابت؟ وهل كل رقم يعبّر عن شيء؟

🤖 **ChatGPT:**

ممتاز... دعني أوضحها بتفصيل شديد:

🧱 **ما معنى [٠,٤٥, ٠,٧٢, ٠,١٨, ..., ٠,٩١]؟**

هذه القائمة التي تراها داخل المعقوفتين **هي "المتجه العددي (Vector)** "الخاص بكلمة واحدة فقط، أو أحيانًا "توكن" واحد.

وكل **رقم في هذا المتجه** يعبّر عن **بُعد رياضي معيّن**.

🧠 **تشبيه تقريبي:**

👨‍🦱 الطالب: هل هناك مثال يقرّب لي الصورة؟

🤖 ChatGPT: نعم.

تخيّل أن عندك **خريطة ذهنية**، وكل كلمة هي **نقطة** في هذه الخريطة.

- الكلمات القريبة في المعنى تكون قريبة في الموقع.
- الكلمات المختلفة في المعنى تكون بعيدة عن بعضها.

"المد" و"مدرّب" قد يكونان قريبين.

لكن "المد" و"جميلة" أبعد قليلًا، لأن معناهما يختلف.

🧭 **القانون الحاكم: القرب في الفضاء العددي**

👨‍🦱 الطالب: هل هناك قاعدة تحكم هذه المسافات بين الكلمات؟

🤖 ChatGPT: نعم، تُسمّى:

قانون القرب(Cosine Similarity)

أو ما يعرف باسم :**القرب في الفضاء المتجهي**

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

كلما اقترب متجهان من بعضهما، كان معناهما متقاربًا.

👨‍🎓 الطالب: ما هو الرقم ٧٦٨ إذًا؟

ChatGPT: 🤖 العدد **٧٦٨** هو:

- 🔸 عدد الأبعاد التي يتكوّن منها المتجه العددي للكلمة.

بمعنى آخر:

- هذه الكلمة تحوّلت إلى متجه طوله **٧٦٨ خانة**.
- أي أن المتجه عبارة عن **٧٦٨ رقمًا متتالٍ**، يمثلون جوانب مختلفة من المعنى.

🎯 تشبيه بسيط:

تخيّل أن عندك إنسانًا، وتريد أن تكتب عنه ملفًا تعريفيًا:

الطول، الوزن، لون العين، نسبة الذكاء، نوع الصوت، قوة التركيز...

حتى تصل إلى ٧٦٨ صفة من صفاته.

نفس الشيء تفعله الآلة.

👨‍🎓 الطالب: وهل كل نموذج يستخدم ٧٦٨ رقمًا؟

ChatGPT: 🤖 لا، هذا **يتغيّر بحسب نوع النموذج**.

النموذج	حجم المتجه (الأبعاد)
Word٢Vec	غالبًا ٣٠٠ بعد فقط
BERT (base)	٧٦٨بعدًا
BERT (large)	١٠٢٤بعدًا
GPT-٣.٥ / GPT-٤	من ١٠٢٤ إلى ٤٠٩٦
GPT-٤٥ (مضغوط ذكي)	متجه ديناميكي — قد يُضغط أحيانًا

🧠 ماذا يفعل النموذج بهذه الأبعاد؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

كل رقم في المتجه يُعبّر عن شيء مثل:

- هل الكلمة تُستخدم في المشاعر؟
- هل هي مجردة أم محسوسة؟
- هل تتكرر في السياق العلمي؟
- هل تظهر بجانب أسماء بشرية؟
- هل تُوصف بها الأماكن؟

...وهكذا.

كلما زادت الأبعاد، كان **الفهم أدق**، لكنه أيضًا **أثقل حسابيًا**.

🧭 **جدول موجز:**

المفهوم	الشرح المبسّط
التوكن	مقطع صغير من الكلمة (مثل: "المد")
التضمين	العملية التي تُحوّل التوكن إلى متجه عددي
المتجه العددي	قائمة أرقام تمثل الكلمة رياضيًّا
الفضاء العددي	الخريطة الرياضية التي توضع فيها المتجهات
قانون القرب	كلما اقتربت المتجهات = زاد التشابه في المعنى
٧٦٨رقمًا	عدد الأبعاد في المتجه الواحد للكلمة في نموذج BERT/GPT

👨‍🎓 الطالب: إذًا، المتجه العددي هو مثل بطاقة هوية ضخمة للكلمة، فيها ٧٦٨ معلومة مشفّرة؟

ChatGPT: 🤖 أحسنت تمامًا.

وهذه **الهوية العددية** هي التي تجعلني أستنتج المعاني، وأربط بين الكلمات، وأفهم السياق،

دون أن أملك فهمًا بشريًا... فقط عبر **الرياضيات**.

📊 **تشبيه بسيط:**

ChatGPT: 🤖 تخيّل أن عندك **بطاقة هوية ضخمة** للكلمة،

كل سطر فيها يصف جانبًا منها:

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

البُعد	القيمة
مدى استعمال الكلمة مع البشر	٠.٨٧
هل تدل على مكان؟	٠.٦٥
هل تُستخدم في التعليم؟	٠.٩٢
هل تُستخدم مع الزمن؟	٠.١٠
هل تُوصف بأنها إيجابية؟	٠.٥٤
...	...

هذه القائمة في الحقيقة هي **المتجه العددي للكلمة**.

👨‍🦱 الطالب: وهل هذا ما يجعل "المدرسة" قريبة من "المعلم" و"الدرس"، وبعيدة عن "الشاحنة" أو "المريخ"؟

ChatGPT: 🤖 أحسنت بالضبط!

لأن متجه "المدرسة" سيقترب كثيرًا من متجه "التعليم" أو "الصف"،

ويبتعد عن متجه "المجرة" أو "البترول..."

🧭 **ملخص مبسّط:**

المفهوم	الشرح المبسط
المتجه العددي(Vector)	قائمة طويلة من الأرقام تمثّل الكلمة بطريقة رياضية
عدد الأرقام	يختلف حسب النموذج (٣٠٠ – ٧٦٨ – ١٠٢٤ – أكثر)
وظيفة كل رقم	يُمثّل بُعدًا لغويًّا أو معنويًّا في استخدام الكلمة
لماذا عددها كبير؟	لتمثيل المعنى الكامل والدقيق للكلمة في مئات الأبعاد

👨‍🦱 الطالب: الآن فهمت!

كلما زادت الأرقام، زادت قدرتك على "رؤية" أبعاد الكلمة المخفية؟

ChatGPT: 🤖 بالضبط.

وهكذا أستطيع أن "أفهم" بدون أن أملك وعياً بشرياً،

فقط عبر **الأرقام الدقيقة** التي تعلّمتها من ملايين الجمل.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

💬 كيف تختلف النماذج في تمثيل الكلمة رقميًّا؟

👨‍🎓 الطالب: قلتَ لي إن كل كلمة تتحوّل إلى متجه عددي...

لكن هل كل النماذج تحوّلها بالطريقة نفسها؟

ChatGPT: 🤖 لا، أبدًا.

كل نموذج له طريقة مختلفة في حساب المتجه العددي،

وهذا يغيّر كثيرًا من النتائج.

دعني أشرح لك **أربع نماذج مشهورة**:

١–البداية الذكية – Word٢Vec

👨‍🎓 الطالب: ما أول نموذج استخدم التمثيل المتّجه؟

ChatGPT: 🤖 اسمه **Word٢Vec**، وابتكرته Google سنة ٢٠١٣.

☑ فكرته بسيطة لكنها ثورية آنذاك:

"قل لي من تُجالس، أَقُل لك من أنت".

👨‍🎓 الطالب: كيف يعني؟

ChatGPT: 🤖 هذا النموذج لا يفهم الكلمة من معناها مباشرة،

بل يراقب **من تجيء بجواره دائمًا**.

🔍 مثال:

- إذا رأى كثيرًا جملًا مثل:

"الطفل يلعب في الحديقة"

"الأطفال يلعبون بالكرة"

"الولد الصغير يلعب يوميًا"

يفهم أن: "يلعب" مرتبطة بـ "الطفل"، "الولد"، "الحديقة..."

📍 النتيجة:

يعطي لكل كلمة متجهًا عدديًا، يمثّل "بيئتها اللغوية."

❌ لكن له عيب:

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الكلمة لها متجه **واحد فقط**،

لا يتغيّر مهما تغيّر السياق.

" 🔁 عين" دائمًا تعني شيئًا واحدًا بالنسبة له... سواء في الطب أو الماء أو الجاسوسية!

٢-رؤية عالمية أوسع – GloVe

👨‍🎓 الطالب: وماذا عنGloVe ؟ سمعت أنه أذكى!

ChatGPT: GloVe 🤖 جاء من Stanford بعدWord٢Vec ،

وهو يُركّز على **الإحصاءات العامة للغة كلها**، ليس فقط الجمل القريبة.

☑فكرته:

لا أراقب جملة واحدة، بل أراقب **كل اللغة**،

كم مرة ظهرت كلمتان معًا في أي مكان!

🔍 مثال:

- إذا ظهرت "قلم" و"مدرسة" كثيرًا في نفس السياقات،

يفهم أن هناك علاقة بينهما.

📌 اسمه اختصار لـ**Global Vectors for Word Representation** :

📍 النتيجة:

- يعطي تمثيلًا عدديًا أفضل لبعض المعاني المجردة
- لكنه لا يزال **ثابتًا** ...الكلمة = متجه واحد فقط!

٣-النموذج السياقي – BERT

👨‍🎓 الطالب: وهل هناك من نماذج أخرى؟

ChatGPT: 🤖 نعم! جاء بعدها نموذج اسمه **BERT**من Google أيضًا.

☑فكرته العبقرية:

الكلمة **لا تُفهم وحدها**، بل **ضمن الجملة**.

👨‍🎓 الطالب: يعني نفس الكلمة يمكن أن يكون لها متجهات مختلفة؟

ChatGPT: 🤖 بالضبط.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

🔍 مثال:

- جملة ١: "أصاب أحمد ألم في **العين**".
- جملة ٢: "شرب من **عين** الماء".
- جملة ٣: "أرسل الحاكم **عينًا** إلى القلعة".

Word٢Vec ❌ يقول: "عين = عين"، دائمًا.

☑ لكن BERT يعطي:

- متجه مختلف لـ "العين" في كل جملة.
- لأنه **يعرف أن معناها تغيّر بحسب السياق**.

📌 اسمه اختصار لـ **Bidirectional Encoder Representations from Transformers**:

٤-نموذج الفهم التوليدي – GPT

👨‍🎓 الطالب: وماذا عنك أنت؟

ChatGPT: 🤖 أنا أستخدم أفكار BERT ، لكن أضيف عليها الكثير.

☑ فكرتي الأساسية:

لا أُمثّل الكلمة فقط... بل أُتوقّع الكلمة القادمة بناءً على كل ما سبق.

👨‍🎓 الطالب: كيف يعني؟

ChatGPT: 🤖 كل كلمة عندي لها متجه خاص بها،

لكن هذا المتجه يتغيّر حسب:

- الجملة السابقة
- هدف المتحدّث
- أسلوب الكتابة
- وحتى نوع المستخدم!

📍 ولهذا نقول إن GPT:

- **ديناميكي**
- **سياقي بعمق**

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

- **قادر على التكيّف حتى مع المعنى الخفي**

📊 جدول مقارنة بين النماذج

النموذج	نوع التمثيل المتّجه	هل يتغيّر حسب السياق؟	مميزاته	عيوبه
Word۲Vec	متجه ثابت	❌ لا	بسيط، سريع، يكتشف العلاقات القريبة	لا يفرّق بين المعاني المتعددة
GloVe	متجه ثابت عالمي	❌ لا	قوي في الإحصاءات الضخمة	لا يفهم السياق
BERT	متجه سياقي	☑ نعم	يفهم المعنى من الجملة نفسها	لا يولّد كلامًا بنفسه
GPT	متجه سياقي ديناميكي	☑ نعم وبعمق	يتوقّع – يولّد – يفهم النبرة والمعنى العميق	يحتاج قوة حسابية أكبر

👨‍🎓 الطالب: يعني لو قلت: "كتبت بقلمي" و"رأيت قلمي في الحلم"،
أنت ستفهم أن "قلمي" هنا واقعي، وهناك رمزي؟

ChatGPT: 🤖 بالضبط! لأنني لا أستخدم متجهًا واحدًا جاهزًا،
بل **أبني المتجه المناسب حسب المعنى والسياق الذي أنت فيه.**

📊 ملخص:

المفهوم	التوضيح
التمثيل المتجه	تحويل الكلمة إلى سلسلة من الأرقام
فضاء المتجهات	فضاء رياضي عالي الأبعاد تتوزع فيه الكلمات
تقارب المتجهات	كلما كانت الكلمات متقاربة في المعنى، تقاربت متجهاتها
التمثيل السياقي	تمثيل الكلمة يتغير بحسب السياق الذي ترد فيه

📌 خاتمة الفصل:

التمثيل المتجه هو الحجر الأساس الذي بُنيت عليه كل تقنيات فهم اللغة في الذكاء الاصطناعي.
بدون هذه الخطوة، لا يمكن للآلة أن تدخل عالم الكلمات.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

وفي الفصل التالي، سنتعرف على مرحلة أخرى:
كيف تُحسب هذه الكلمات؟ وما المقصود بـ "التوكنات"؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل الثاني: التوكنات والتكلفة؟

📌 توطئة الفصل:

بعد أن فهمنا كيف تتحوّل الكلمات إلى متجهات، نحتاج إلى معرفة كيف تتعامل النماذج مع النصوص على المستوى الأدنى: التوكنات. ما هي التوكنات؟ ولماذا لا تُحسب الكلمات كما هي؟ وما علاقة ذلك بتكلفة الاستخدام؟

👨‍🎓 الطالب: ما المقصود بكلمة "توكن"؟

🤖 ChatGPT: التوكن (Token) هو وحدة صغيرة من النص يستخدمها النموذج في المعالجة. قد تكون الكلمة كاملة، أو جزءًا منها، أو حتى رمزًا أو علامة ترقيم.

ولو كانت الكلمة الواحدة تنقسم إلى **توكنين أو ثلاثة**، فالتكلفة تزداد.

👨‍🎓 الطالب: لماذا لا يتعامل النموذج مع الكلمات مباشرة؟

🤖 ChatGPT: لأن الكلمات تختلف كثيرًا من حيث الطول والاستخدام. فبدلًا من معالجة كل كلمة كعنصر مستقل، يتم تقطيع النص إلى وحدات أصغر تُسمى توكنات. هذا يجعل المعالجة أكثر كفاءة ومرونة.

👨‍🎓 الطالب: هل يمكنك إعطائي مثالًا؟

🤖 ChatGPT: بالتأكيد. كلمة مثل: "المهندس" قد تتحول إلى توكنين:

- "ال" "مهندس"
- "أما كلمة مثل: "في" فقد تبقى توكنًا واحدًا.

🧩 مثال توضيحي:

الجملة	عدد التوكنات
"المدرسة جميلة"	٣ توكنات: "المد" + "رسة" + "جميلة"
"معلومات دقيقة"	٤ توكنات تقريبًا
"الذكاء الاصطناعي"	قد تكون ٥ أو ٦ توكنات

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

كلما كانت الكلمات أكثر **تركيبًا أو تعقيدًا**، زاد عدد التوكنات.

👨‍🎓 الطالب: وما علاقة هذا بالتكلفة؟

ChatGPT: 🤖 تكلفة استخدام النموذج – خاصة في الخدمات السحابية مثل OpenAI – تُحسب بناءً على عدد التوكنات التي تُستخدم في الإدخال والإخراج.

كلما زاد عدد التوكنات، زادت التكلفة.

👨‍🎓 الطالب: هل التوكن يساوي كلمة؟

ChatGPT: 🤖 لا دائمًا. في اللغة العربية، قد تحتوي الكلمة على لاحقة أو بادئة، فتنقسم إلى أكثر من توكن. بينما في كلمات إنجليزية قصيرة قد تبقى الكلمة توكنًا واحدًا.

🧪 **مثال:**

جملة "أنا أدرس في جامعة الإمام محمد بن سعود" قد تتكوّن من ٩–١٢ رمزًا.

يمكن الحصول على هذا التطبيق من هنا: انقر على الرابط

📍 **تطبيق:**

في اختبار لغوي رقمي، يُحسب عدد الرموز وليس عدد الكلمات، مما يؤثر على طول الذاكرة، ويؤثر على الطول الذي يستطيع النموذج معالجته.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

تطبيق:

١. عند استخدام ChatGPT للواجبات المدرسية، قد يظهر لك في الأدوات المتقدمة أن الحد الأقصى هو "٤٠٩٦"Token ، وليس "٤٠٩٦ كلمة."

٢. هذا يعني أن مقالًا يحتوي على ١٠٠٠ كلمة قد يُرفض لأنه فعليًّا يعادل ٣٠٠٠ – ٣٥٠٠ Token.

👨‍🎓 الطالب: لقد فهمتُ أن النص يُحوّل إلى رموز...

لكن هل لهذه الرموز علاقة بالتكلفة المالية؟

🤖 ChatGPT: نعم، تمامًا.

الرمز (Token) هو **وحدة القياس الأساسية** التي يُحسب بها الجهد الذي أبذله في كل تفاعل.

وهي أيضًا **الأساس الذي تُحدَّد عليه التكلفة المالية لاستخدام النموذج**.

👨‍🎓 الطالب: وكيف ذلك؟ وضّحها لي خطوة خطوة.

🤖 ChatGPT: حين ترسل إليّ سؤالًا، أقوم بثلاث عمليات:

١. أقرأ الرموز التي كتبتها (مدخلات)

٢. أعالجها داخليًا (تفكير)

٣. وأردّ عليك برموز جديدة (مخرجات)

كل رمز أقرؤه أو أكتبه، يدخل في **حساب الجهد والتكلفة**.

👨‍🎓 الطالب: وهل معرفة عدد الـ Tokens تهمني فعلًا كمستخدم؟

🤖 ChatGPT: بالتأكيد، لأنها تؤثر في:

١. طول النص الذي يمكنك إرساله دفعة واحدة.

٢. مدى تذكّري لما قلتَه في البداية (يتعلّق بالسياق).

٣. جودة الاستجابة، كلما كانت الرسالة قصيرة ومباشرة، كانت الإجابة أفضل.

٤. في بعض الخدمات المدفوعة، يُحسب السعر بناءً على عدد الرموز.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

📝 جدول توضيحي: ما الذي يُعدToken ؟

العنصر	هل هوToken ؟	ملاحظات
كلمة (مثل: "مدرسة")	☑ نعم	إذا كانت ضمن القاموس
جزء من كلمة (مثل: "ال" + "مدرسة")	☑ نعم	يُجزّأ إن لم يكن محفوظًا ككلمة واحدة
علامة ترقيم (؟ ، .)	☑ نعم	رمز مستقل
مسافة بين الكلمات	❌ لا	لا تُعدToken

🧭 ملخص معرفي:

المصطلح	الشرح المبسط
التوكن	مقطع صغير من الكلمة (مثل: "المد" – "رسة")
لماذا يُستخدم؟	لتكراره في كلمات كثيرة + لتخزين أفضل
هل هو مقروء؟	أحيانًا لا، لكنه مفيد للنموذج حسابيًّا
هل يفيد في الفهم؟	نعم، لأنه يساعد في التعميم ومعالجة كلمات جديدة

🧪 مقارنة توضيحية:

الكلمة الأصلية	التوكنات الناتجة	ملاحظات التعميم
المدرسة	"المد" + "رسة"	"المد" واضح (مدّ البحر، مدّ الصوت)، "رسة" غير شائع
المعلومات	"المعلو" + "مات"	كلاهما يُستخدم في مفردات أخرى ⟵ تعميم أوضح وأقوى

☑ كلمة بديلة أخرى:

"الكتابة"

تقسيم "الكتابة" ⟶ "الكتا" + "بة"

التوكن	كلمات أخرى مشابهة
"الكتا"	الكتاتيب – الكتّاب – الكاتب – مكتبة
"بة"	لعبة – هبة – كتبة – خطبة

- "المد" = توكن ١
- "رسة" = توكن ٢

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

- "جميلة" = توكن ٣

وهذه هي المدخلات الخام التي أبدأ بها التحليل.

👨‍🎓 الطالب: وهل الرموز مجرد أجزاء من الكلام... أو هي أيضًا وحدات محاسبة؟

ChatGPT: 🤖 بالضبط.

كما تُقاس الكهرباء بالكيلوواط، أو الإنترنت بالميغابايت،

أنا أقيس بعدد الرموز التي أتعامل معها.

👨‍🎓 الطالب: وهل تُحسب التكلفة على كل رمز؟

ChatGPT: 🤖 لا، لا تُحسب على كل رمز بمفرده...

بل تُحسب عادة **لكل مليون رمز/ توكنات**،

وهو مقياس عالمي معتمد لتحديد التكلفة.

يتم التفريق بين:

- 🔹 **المدخلات** :الرموز التي ترسلها أنت إليّ(Input)
- 🔹 **المخرجات** :الرموز التي أرسلها أنا إليك(Output)

وكل واحدة منهما تُحسب بتعريفة مختلفة.

🧪 **مثال توضيحي:**

إذا استخدمت نموذج GPT-٤o ، فستكون التكلفة:

النوع	السعر لكل مليون توكنات
المدخلات/ السؤال	٥.٠٠دولار
المخرجات/ الرد	١٥.٠٠دولار

يعني أنك لو أرسلت لي نصوصًا مجموعها مليون رمز، ستدفع ٥ دولارات للمدخلات،

ولو طلبت مني ردودًا بحجم مليون رمز، ستدفع ١٥ دولارًا للمخرجات.

📍 **تطبيق تربوي:**

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

في منصة تعليمية تستخدم GPT للإجابة على استفسارات الطلاب:

تُقدّر الكلفة الشهرية بحسب إجمالي الرموز المستخدمة في الأسئلة والأجوبة، مثلًا:

(٣ ميلون) مدخل –الإجمالي = ١٥ دولار

(٣ مليون) مخـرج –الإجمالي = ٤٥ دولار

👨‍🎓 الطالب: فهل يعني هذا أن الاستخدام الطويل مكلف؟

ChatGPT: 🤖 يعتمد على النموذج والسعة.

لكن الفهم الجيد لـ **عدد الرموز = يعطيك تحكمًا أفضل في التكلفة والجودة.**

📍 **تطبيق واقعي:**

- أرسلت وثيقة بحثية ضخمة: ١١٠،٠٠٠ Token

⟵النموذج يستطيع الرد بـ ١٨،٠٠٠ فقط

⟵وقد لا يستطيع التلخيص الكامل إن كان النص يحتاج مساحة ردّ كبيرة.

- أرسلت ٥٠،٠٠٠ Token فقط:

⟵سيتبقى ٧٨،٠٠٠ Token للرد

⟵يمكنني إنتاج رد غني، تلخيص، تحليل، وحتى تقسيمات هيكلية.

👨‍🎓 الطالب: كيف أضمن استخدامًا ذكيًا للرموز إذًا؟

ChatGPT: 🤖 إليك ثلاث قواعد ذهبية:

١. لا ترسل الحد الأقصى كاملًا (١٢٨K) إذا كنت تريد ردًّا.

٢. قسّم النصوص الكبيرة إلى أجزاء صغيرة.

٣. راقب حجم المخرجات المتوقعة: هل تريد ملخصًا؟ تحليلًا؟ كلما كبر، قلّل المدخلات.

👨‍🎓 الطالب: ممتاز... هكذا أفهم التكلفة قبل أن أبدأ.

ChatGPT: 🤖 وهذا هو الذكاء في استخدام الذكاء الاصطناعي!

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

🧪 **مثال عملي:**

👨‍🎓 الطالب: سمعت أن GPT-٤٥ يستطيع قراءة ١٢٨ ألف رمز... ماذا يعني هذا بالضبط؟

ChatGPT: 🤖 يعني أنني أستطيع أن أُبقي في ذاكرتي اللحظية حتى **١٢٨,٠٠٠** توكنات من النص الذي ترسله إليّ دفعة واحدة.

👨‍🎓 الطالب: وهل هذا كثير؟

ChatGPT: 🤖 نعم، هذا ضخم جدًّا مقارنة بالنماذج السابقة.

GPT-٣.٥= كانت نافذته ٤،٠٠٠ Token فقط

GPT-٤= الأساسي كانت ٨،٠٠٠ أو ٣٢،٠٠٠Token

أما GPT-٤٥ = يستوعب **١٢٨,٠٠٠ Token**،

أي حوالي:

١٠٠,٠٠٠ 📝 **كلمة** = **٣٥٠** 📄 **صفحة– A٤ تقريبًا** (حسب اللغة والتنسيق)

👨‍🎓 الطالب: ماذا أستفيد من هذه السعة الكبيرة؟

ChatGPT: 🤖 فوائد كثيرة، أهمها:

١. **تلخيص كتب كاملة** دفعة واحدة دون تجزئة

٢. **قراءة محاضر أو بحوث طويلة** دون أن أنسى البداية

٣. **إجراء تحليل شامل لوثائق قانونية أو طبية أو تعليمية** في جلسة واحدة

٤. **مواصلة الحوار العميق دون فقدان الترابط**

🧪 **مثال عملي:**

أرسلت إليّ ملفًا يحتوي على:

- ملخصات ١٠ فصول دراسية
- أوراق عمل
- ملحق مراجع

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

⟵إن كان مجموعها في حدود ١٢٠،٠٠٠ توكنات

⟵أستطيع أن أقرأها كلها **دفعة واحدة**

⟵وأجيبك بإجابة تأخذ جميع السياقات السابقة بالحسبان

📍 **تطبيق تربوي:**

- أستاذ جامعي يرفع خطة مقرر + أسئلة اختبار + مشاريع الطلبة،

ثم يسأل النموذج عن أوجه الضعف الشائعة ⟵ النموذج يتذكر كل شيء لأن كل المحتوى في نافذته!

- طالب يرفع كتاب "البلاغة العربية" ويريد تلخيصه مع تحديد مواضع الاستطراد ⟵ لا حاجة لتقسيمه!

👨 الطالب: وهل استخدام هذه النافذة له تكلفة أيضًا؟

ChatGPT: 🤖 نعم، كلما زاد عدد الرموز المستخدمة داخل النافذة،

زادت **تكلفة المعالجة** بحسب عدد الرموز الداخلة في السياق،

حتى لو لم تكن كلها "نصًا ظاهرًا."

📜 **جدول توضيحي:**

النافذة	عدد الرموز	ما يعادلها من الكلمات	الاستخدامات
GPT-٣.٥	٤٠٠٠ توكنات	٣٠٠٠ كلمة	ردود سريعة – محادثة
GPT-٤	٣٢٠٠٠ توكنات	٢٤٠٠٠ كلمة	مقالات – فصول مفردة
GPT-٤o	١٢٨٠٠٠ توكنات	١٠٠ ألف كلمة	كتب كاملة – تحليلات معمقة

👨‍🎓 الطالب: وهل هذه الذاكرة دائمة؟

ChatGPT: 🤖 لا.

هذه **ذاكرة مؤقتة داخل الجلسة الواحدة فقط،**

بمجرد انتهائها أو إغلاق المحادثة، تُفقد المعلومات...

هل ١٢٨,٠٠٠ Token تعني أن بإمكاني إرسال ١٢٨ ألف فقط؟ أم تشمل الرد أيضًا؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

👨 الطالب: فهمتُ أن GPT-٤٥ يقرأ حتى ١٢٨ ألفToken ،
فهل هذا يعني أنني أستطيع إرسال ملف بحجم ١٢٨ ألف Token دفعة واحدة؟

ChatGPT: 🤖 سؤال مهم جدًا.
في الواقع: **لا يمكنك أن ترسل ١٢٨ ألف وتنتظر أن أردّ بعد ذلك،**
لأن هذا الحد يشمل **المجموع الكلي للمدخلات والمخرجات معًا**.

👨 الطالب: إذًا لو أرسلت لك ملفًا حجمه ١٢٨ ألف Token بالضبط، ماذا يحصل؟
ChatGPT: 🤖 سأتلقّاه...
لكنني **لن أستطيع الردّ بأي رمز واحد.**
لأنني وصلت إلى السعة القصوى، ولن يتبقى مجال للإجابة.

🧪 **مثال عملي:**

حجم المدخلات	المتاح للمخرجات	هل يمكن الرد؟
١٠٠،٠٠٠ Token	٢٨،٠٠٠ Token	☑ نعم
١٢٨،٠٠٠ Token	٠ Token	❌ لا
٦٤،٠٠٠ Token	٦٤،٠٠٠ Token	☑ نعم

👨 الطالب: وما علاقة هذا بالتكلفة المالية؟
ChatGPT: 🤖 علاقة مباشرة جدًا:

النوع	يُحسب عليه تكلفة؟
ما ترسله لي	☑ نعم (مدخلات)
ما أرسله لك	☑ نعم (مخرجات)

فلو أرسلت ٨٠،٠٠٠ Token وقلت لك: "لخّص هذا"،
ثم رددت عليك بـ ٢٠،٠٠٠Token ،
⟵تُحسب التكلفة على إجمالي ١٠٠,٠٠٠ Token

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

ملخص:

المفهوم	التوضيح
التوكن	وحدة صغيرة من النص تُستخدم في التحليل
أداة التقطيع	الأداة التي تُحوّل النص إلى توكنات
تكلفة التوكن	عدد التوكنات المستعملة في الإدخال والإخراج يُحدد السعر
التقطيع الجزئي	الكلمة قد تُقسم إلى أكثر من توكن

خاتمة الكتاب:

فهم التوكنات مهم ليس فقط لتحليل كيفية معالجة النموذج للنص، بل أيضًا لفهم التكلفة، وتحسين الاستخدام.

وفي الفصل التالي، ننتقل من الوحدات الصغيرة إلى سياق الكلمة: كيف تُفهم الكلمة بناء على الجملة؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل الثالث: التضمين السياقي: كيف تفهم الآلة اختلاف المعنى؟

📌 توطئة الفصل:

ليست كل الكلمات تُستخدم بنفس المعنى في كل المواضع. كلمة مثل "عين" في "عين الماء" غيرها في "عين الجاسوس". فهل تفرّق الآلة بينهما؟ في هذا الفصل، نستكشف كيف تلتقط النماذج الحديثة التغير في المعنى بحسب السياق.

👨‍🎓 الطالب: ما معنى "التضمين السياقي"؟

🤖 ChatGPT: هو قدرة النموذج على تغيير تمثيل الكلمة بحسب السياق الذي وردت فيه.

الكلمة لا يكون لها متجه واحد ثابت كما في النماذج القديمة **مثل:Word۲Vec** ،

بل يتغيّر حسب الجملة

👨‍🎓 الطالب: هل هذا يختلف عنWord۲Vec ؟

🤖 ChatGPT: نعم Word۲Vec .يعطي الكلمة تمثيلًا ثابتًا، لا يتغيّر مهما اختلف السياق.

أما النماذج الحديثة مثل BERT وGPT، فهي تنتج تمثيلًا سياقيًّا،

أي أن متجه الكلمة يتشكّل أثناء قراءتها داخل الجملة.

👨‍🎓 الطالب: هل يمكنك إعطائي مثالًا للنماذج القديمة؟

🤖 ChatGPT: بالتأكيد. الكلمة "عين" في:

- "عين إنسان"
- "عين ماء"
- "عين جاسوس"

"عين" في كل الجمل السابقة تحمل نفس الأرقام وهذا يجعل الفهم محدودًا وغير دقيق.

👨‍🎓 الطالب: وماذا عن النماذج الحديثة؟

🤖 ChatGPT: الوضع هنا يختلف، فكلمة "عين" يتم تحديد معناها بناء على السياق،

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

وسيتم تمثيلها بمتجهين مختلفين داخل النموذج، لأن الكلمات المحيطة بها مختلفة.

وتستطيع الآلة أن "تفهم" المعنى **بالأرقام**،

فتصبح المعاني على النحو الآتي:

- في جملة: "ألم في العين" ⟵ متجه يشير إلى عضو بصري
- في: "شرب من عين ماء" ⟵ متجه يشير إلى نبع
- في: "أرسل الحاكم عينًا" ⟵ متجه يشير إلى جاسوس

نفس الكلمة، لكن ثلاث معاني = ثلاث متجهات مختلفة!

👨‍🎓 الطالب: وماذا نقول في كلمة "جميلة" ألا تظل تحمل معنى "جميلة"؟

ChatGPT: 🤖 سؤال وجيه، وللرد عليه تنبه إلى هذا الاختلاف بين الجمل:

- "المدرسة جميلة" ⟵ المتجه يشير إلى الجمال المكاني أو الهندسي.
- "الروح جميلة" ⟵ المتجه يشير إلى الجمال الأخلاقي أو المعنوي.
- "الخطابة جميلة" ⟵ المتجه يشير إلى الجمال البلاغي أو التعبيري.

كل سياق ينتج **متجهًا مختلفًا لنفس الكلمة**.

👨‍🎓 الطالب: ما فائدة هذا التضمين السياقي؟

ChatGPT: 🤖 يجعل النموذج أكثر فهمًا ودقة في مهام مثل: الترجمة، الإجابة على الأسئلة، تلخيص النصوص، لأن التمثيل يصبح أقرب إلى الطريقة التي نفهم بها نحن البشر.

📌 خاتمة الفصل:

التضمين السياقي يمثل قفزة نوعية في طريقة فهم الآلة للكلمات. لقد انتقلنا من المعنى الثابت إلى الفهم المرن، ومن الحفظ إلى السياق. وفي الفصل التالي، سنتعمق أكثر: كيف يختار النموذج كلمة دون أخرى؟ وما سر التنبؤ الإحصائي؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل الرابع: التنبؤ الإحصائي: لماذا يختار النموذج كلمة دون أخرى؟

📌 توطئة الفصل:

حين نكتب جملة للنموذج، فإنه لا يختار الكلمات عشوائيًا، بل يُجري عمليات حسابية معقدة ليحدّد الكلمة الأكثر احتمالًا أن تكون التالية. في هذا الفصل نكشف آلية التنبؤ، وكيف تعمل النماذج كمولّدات احتمالية.

👨‍🎓 الطالب: ماذا يعني أن النموذج "يتنبأ" بالكلمة؟

🤖 ChatGPT: عندما تعطي النموذج جملة ناقصة،

فإنه يحسب احتمالات الكلمات الممكنة التالية،

ويختار أعلى احتمال.

فالتنبؤ يتم على أساس إحصائي مبني على البيانات التي تدرب عليها.

👨‍🎓 الطالب: ما شكل هذا التنبؤ؟

🤖 ChatGPT: النموذج لا يرى الكلمات، بل يرى تمثيلاتها الرقمية.

ثم يحسب احتمالات، مثل:

- "مدرسة" ⟶ ٠.١٢
- "جامعة" ⟶ ٠.٣٥
- "مسجد" ⟶ ٠.٠٧ "ويختار "جامعة" لأنها صاحبة أعلى احتمال.

👨‍🎓 الطالب: هل يعتمد هذا على السياق؟

🤖 ChatGPT: بالتأكيد. لو كتبت: "بعد الثانوية العامة التحق بـ"، فالنموذج يعرف من السياق أن "جامعة" أكثر منطقية من "مدرسة"، وهكذا.

👨‍🎓 الطالب: هل هناك خوارزمية معينة يستخدمها؟

🤖 ChatGPT: نعم. هناك ما يُعرف بـSoftmax ، وهي دالة تحويل تحوّل الدرجات العددية إلى احتمالات. كما تُستخدم شبكات عصبونية لتقييم الخيارات بناءً على الوزن والتعلم السابق.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

👨‍🎓 الطالب: إذن النموذج لا يملك نية، بل مجرد حساب؟

🤖 ChatGPT: بالضبط. هو لا يفهم كما يفهم الإنسان، لكنه بارع في رصد الأنماط، واختيار ما ينسجم مع ما سبق.

🧪 **مثال تراكمي:**

لو كتبت:

"كان الجو..."

سأكملها بناءً على تدريبي، وقد أقدّر:

- "ممطرًا" = ٤٠ ٪
- "جميلاً" = ٢٨ ٪
- "باردًا" = ١٥ ٪
- "مظلمًا" = ٦ ٪
- "غائمًا" = ٥ ٪

⟵أختار "ممطرًا"، ثم أعيد الحساب للجملة الجديدة:

"كان الجو ممطرًا"...

ثم أكمل بناءً على الاحتمالات التالية.

📊 ملخص:

المفهوم	التوضيح
التنبؤ الإحصائي	اختيار الكلمة التالية بناءً على أعلى احتمال إحصائي
دالةSoftmax	دالة تحول الأوزان إلى احتمالات
السياق	ما يسبق الكلمة يؤثر في قرار اختيارها
غياب الفهم الحقيقي	النموذج لا يفهم المعنى، بل يبني على الأنماط والتكرار

📌 خاتمة الفصل:

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

نحن الآن أمام نموذج يتصرف كما لو كان كاتبًا، لكنه في الحقيقة مجرد آلة احتمالات. التنبؤ الإحصائي هو القلب النابض لنماذج اللغة، ومن هنا ننتقل إلى سؤال أعمق: كيف تُقرّر هذه النماذج ما هو الأهم في الجملة؟ الجواب في الفصل القادم عن "الانتباه الذاتي."

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

🧠 الفصل الخامس: الانتباه الذاتي: كيف يعرف النموذج ما المهم في الجملة؟

📌 توطئة الفصل:

حينما تقرأ جملة ما، قد تقفز عيناك إلى الكلمات الأساسية التي تحوي المعنى، وتتجاهل الزوائد. هذه القدرة على التركيز ليست حكرًا على الإنسان، بل تعلّمتها النماذج اللغوية من خلال ما يُعرف بـ "الانتباه الذاتي". إنها الآلية التي تجعل النموذج يركّز على الكلمات المهمة في كل جملة.

لا يكفي أن يعرف النموذج معنى الكلمات، بل عليه أن يعرف أيها أهم من الأخرى. في هذا الفصل نستعرض آلية تُعدّ من أذكى أدوات النماذج الحديثة: "الانتباه الذاتي"، وكيف تجعل النموذج يركّز على ما يهم داخل الجملة.

👨‍🎓 الطالب: ما هو "الانتباه الذاتي"؟

🤖 ChatGPT: هو آلية تُستخدم داخل النماذج لتحديد الأجزاء الأهم في الجملة، من خلال حساب وزن لكل كلمة بحسب علاقتها ببقية الكلمات. النموذج لا يعامل كل الكلمات بالتساوي.

👨‍🎓 الطالب: كيف يفعل ذلك؟

🤖 ChatGPT : هو آلية تجعلني – داخل كل طبقة – أُعطي **وزنًا مختلفًا** لكل كلمة بناءً على علاقتها بالكلمات الأخرى في الجملة.

الكلمات الأكثر أهمية تُعطى وزنًا أعلى،

والكلمات الأقل أهمية تُعطى وزنًا أقل،

كل كلمة تُحوَّل إلى تمثيل عددي،

ثم يقوم النموذج بحساب درجة الأهمية النسبية لكل كلمة بالنسبة لبقية الكلمات في الجملة.

👨‍🎓 الطالب: كيف يعرف النموذج ما المعنى المقصود؟

🤖 ChatGPT: النموذج يستخدم آلية الانتباه الذاتي Self-Attention ، ليرى الكلمات التي حول الكلمة المستهدفة، ويُعطيها وزنًا بحسب قربها وأهميتها في تحديد المعنى.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

👨 الطالب: كيف يعرف أي كلمة أهم؟

ChatGPT: 🤖 عندما يقرأ النموذج جملة مثل "الطالب كتب الدرس بعناية"، فإنه يُعطي كلمات مثل "كتب" و"بعناية" وزنًا أعلى لأنها تحمل المعنى الأساسي، بينما قد يكون وزن "ال" و"الدرس" أقل.

👨 الطالب: هل يمكن ضرب مثال؟

ChatGPT : 🤖 إذا قلت: "ذهب أحمد إلى المكتبة لأنه يحب القراءة"، فإن كلمة "لأنه" ستجعل النموذج يعطي أهمية أكبر لـ "يحب" و"القراءة" لفهم سبب الذهاب.

👨 الطالب: هل هذا يشبه تسليط ضوء على الكلمات الأهم؟

ChatGPT : 🤖 بالضبط. تخيّله كمصباح يدور بين الكلمات، ويُضيء بدرجات مختلفة بناءً على أهمية كل كلمة.

👨 الطالب: وما فائدة هذا الانتباه؟

ChatGPT : 🤖 يجعل النموذج أكثر فهمًا للسياق، فيستطيع الترجمة، والتلخيص، والإجابة بدقة أعلى.

👨 الطالب: هل هذه الآلية موجودة في كل نماذج اللغة؟

ChatGPT : 🤖 ظهرت بشكل جزئي في نماذج سابقة، لكن استخدامها الكامل بدأ مع نموذج "Transformer"، الذي بنيت عليه نماذج مثل GPT و BERTوغيرها.

👨 الطالب: وهل يُحسب هذا الوزن تلقائيًا؟

ChatGPT: 🤖 نعم. النموذج يُجري عمليات رياضية تسمىAttention Scores ، وتُحسب باستخدام المتجهات الناتجة عن الكلمات. ثم يُطبّق هذه الأوزان أثناء توليد التمثيل النهائي للجملة.

👨 الطالب: وهل الانتباه الذاتي هو نفسه في كل النماذج؟

ChatGPT: 🤖 لا. بعض النماذج تستخدم Attention أحادي الاتجاه مثلGPT ، وبعضها ثنائي الاتجاه مثل BERT ، وبعضها يعتمد على توزيع الأوزان فقط، وبعضها يدمجها مع آليات أخرى.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

👨‍🎓 الطالب: هل يمكن تشبيه هذا الانتباه؟

ChatGPT: 🤖 نعم. تخيّل أنك تقرأ مقالًا وتُبرز بقلمك بعض الجمل. هذه الجمل هي ما انتبهت إليه أكثر. النموذج يفعل شيئًا شبيهًا، لكنه بلغة الرياضيات.

🧪 **مثال تطبيقي:**

لو قلت:

"سافر خالد إلى الرياض لأنه يحبها كثيرًا".

حين أصل إلى كلمة "يحبها"،

عليَّ أن أعرف أن "ها" في كلمة (يحبها) تشير إلى "الرياض"، وليس "خالد."

وهنا يقوم نموذج GPT بإعطاء **وزن أعلى لكلمة "الرياض** "لأنها أكثر ارتباطًا بالسياق.

🧪 **مثال:**

لو قلت:

"الطالبة لم تحضر الدرس لأنها كانت مريضة".

عندما أقرأ كلمة **"كانت** "أو **"مريضة"**،

أحتاج أن أعرف: من التي كانت مريضة؟

- "الطالبة" ⟵ المرجع الصحيح
- وليس "الدرس"

وهنا يقوم الانتباه الذاتي بإعطاء **وزن أعلى** لكلمة "الطالبة"،

لأنها ترتبط مباشرة بسياق الضمير "هي."

🧪 **أمثلة وتطبيقات:**

١. **نموذج الانتباه داخل جملة:**

"رأى الطالبُ المعلمَ في الفصل".

النموذج يعطي وزنًا لـ"المعلم" عندما يتوقع كلمة "الفصل"، لأنها ترتبط بها منطقيًا.

٢. **مفاجأة لغوية:**

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

"الكتاب الذي قرأته أمس لم أفهمه لأنه"...

يتوقع النموذج كلمة مثل "معقد" أو "غامض"، لأنه تعلّم أن هذا الأسلوب شائع في اللغة.

٣. **جملة غامضة:**

"دخل محمد على علي وهو يضحك"

يحاول النموذج أن يخمّن من هو الذي "يضحك"، بناءً على السياقات الشائعة.

🧪 **مثال رقمي مبسّط:**

👨‍🎓 الطالب: كيف تفرّق بين كلمتين مثل "مدرسة" و"مستشفى"؟

ChatGPT: 🤖 تمثيلهما العددي مختلف، لكنهما يشتركان في بعض الأبعاد:

- "مدرسة" ترتبط بـ (تعليم، طلاب، دروس، فصول)
- "مستشفى" ترتبط بـ (علاج، مريض، طبيب، صحة)

⟵لكن لو قلت: "ذهبتُ إلى المدرسة للعلاج"، سيظهر خلل إحصائي في التمثيل.

📊 ملخص:

المفهوم	التوضيح
التضمين السياقي	تمثيل الكلمة يتغيّر بحسب السياق
الانتباه الذاتي	آلية لتحديد أهمية كل كلمة في السياق
النماذج السياقية	نماذج مثل BERT و GPTالتي تنتج تمثيلات ديناميكية للكلمات
Attention Scores	أوزان رقمية تُحسب لتحديد تأثير كل كلمة
التوزيع غير المتساوي	بعض الكلمات لها تأثير أعلى من غيرها

📌 خاتمة الفصل:

الانتباه الذاتي جعل النماذج أكثر قربًا لفهم الإنسان. لم تعد الكلمات متساوية، بل صار للنموذج قدرة على تمييز الأهم من المهم. وهذا ما يمهّد للفصل القادم حول البنية الداخلية التي تُنسّق كل هذه العمليات: "معمارية المحوّل". حسب الفصول:

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل السادس: معمارية المحوّل: كيف تُنسّق كل هذه العمليات؟

📌 توطئة الفصل:

الآن وقد فهمنا كيف تنتبه النماذج للكلمات وتُعطيها أوزانًا، حان الوقت لكشف الهيكل الكامل الذي يجمع كل هذه العمليات في منظومة واحدة. نرحّب بك في قلب الذكاء الاصطناعي اللغوي: "المحوّل" أو Transformer.

👨‍🎓 الطالب: ما هو "المحوّل"؟

🤖 ChatGPT: المحوّل هو بنية نموذجية معمارية (Model Architecture) طُورت لمعالجة اللغة بطريقة تسمح بفهم العلاقات بين الكلمات في الجملة كلها دفعة واحدة، باستخدام آلية "الانتباه الذاتي."

👨‍🎓 الطالب: ما الفرق بينه وبين الشبكات التقليدية؟

🤖 ChatGPT: الشبكات التقليدية) مثل (RNN كانت تقرأ الكلمات بالتسلسل، كلمة كلمة. أما المحوّل فينظر إلى الجملة كلها في آن واحد، ويحسب علاقات كل كلمة بكل كلمة أخرى.

👨‍🎓 الطالب: ممّ يتكوّن المحوّل؟

🤖 ChatGPT: يتكوّن من طبقات متكررة، وكل طبقة تحتوي عادةً على:

١. طبقة انتباه ذاتي.
٢. طبقة تغذية أمامية.(Feedforward)
٣. آليات تطبيع وتوصيل) مثل.(LayerNorm

👨‍🎓 الطالب: هل هناك شكل بسيط لفهم هذه البنية؟

🤖 ChatGPT: نعم، تخيّل المحوّل كسلسلة من المحطات، كل محطة تعيد تقييم الكلمات من جديد بحسب سياقها، ثم ترسل النتيجة للمحطة التالية. كلما مرّت الكلمة عبر محطة، زادت دقّتها في التمثيل.

👨‍🎓 الطالب: وهل تختلف هذه البنية في BERT و GPT؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

ChatGPT: 🤖 نعم BERT .يستخدم محوّلًا ثنائي الاتجاه (يرى ما قبل وبعد الكلمة)، أما GPT فيستخدم محوّلًا أحادي الاتجاه (يرى ما قبل فقط). الفرق جوهري في طبيعة الفهم والتوليد.

📊 ملخص المفاهيم:

المفهوم	التوضيح
المحوّل(Transformer)	بنية نموذج تعتمد على الانتباه الذاتي لمعالجة اللغة
طبقات المحوّل	سلسلة من المحطات تمرّ بها الكلمات لإعادة التمثيل
الفرق عنRNN	المحوّل ينظر للجملة دفعة واحدة لا كلمة كلمة
الفرق بين BERT وGPT	BERTثنائي الاتجاه، GPTأحادي الاتجاه

📌 خاتمة الفصل:

معمارية المحوّل هي الإطار العام الذي ينسّق كل ما سبق: من التوكنات، إلى الانتباه، إلى التنبؤ. إنها العمود الفقري للنماذج اللغوية الحديثة، وهي ما سيمهد لنا لفهم الطبقات الخفية القادمة داخل النموذج.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل السابع: الطبقات الخفية: هل له "عقل" متعدد الطبقات؟

📌 توطئة الفصل:

حينما نتحدث عن ذكاء النموذج،

فإننا لا نقصد ذكاءً سطحيًا يعتمد على كلمة بكلمة،

بل نعني قدرة النموذج على فهم مركّب، طبقي، يتعمق في المعنى خطوة بعد خطوة.

هذه القدرة ليست سحرًا،

بل نتيجة بنية داخلية تعتمد على "طبقات خفية"

تعالج المعلومة بمستويات متعددة من التجريد.

وكلما قرأنا أكثر عن النماذج، نجد أنها مكوّنة من "طبقات".

لكن ما المقصود بهذه الطبقات؟ وهل تشبه طبقات التفكير عند الإنسان؟

في هذا الفصل نغوص داخل النموذج لنكشف كيف تُبنى معاني معقّدة من خلال تراكم تمثيلات خفية.

👨‍🦱 الطالب: سمعت بعض الشيء عن "الطبقات الخفية" في الشبكات العصبية، لكن لا أفهم ما المقصود بها تحديدًا. هل يمكنك شرحها لي ببساطة؟

🤖 ChatGPT: بالتأكيد. الطبقات الخفية هي ببساطة وحدات داخلية في النموذج تتلقى المعلومات، وتعالجها، ثم تنقلها إلى الطبقة التي تليها. كل طبقة تتعلّم شيئًا مختلفًا عن النص.

👨‍🎓 الطالب: وهل كل طبقة تفعل الشيء نفسه؟

🤖 ChatGPT: لا. كل طبقة لها وظيفة مختلفة بعض الشيء.

الأولى تفحص البنية اللغوية — ترتيب الكلمات

الثانية تقدّر العلاقات النحوية (فاعل، مفعول، موصوف...)

الثالثة تُكوّن صورة أوّلية عن المعنى العام للجملة

الرابعة تلتقط السياق — هل الجملة تعليمية؟ سؤال؟ عاطفية؟

الأخيرة تصوغ الرد أو التنبؤ التالي، بناء على ما فُهِمَ سابقًا

النتيجة: أنه كلما زادت الطبقات، زادت قدرتي على الفهم الدقيق.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

👨‍🎓 **الطالب: وكيف يتعلّم النموذج كل هذا؟**

🤖 ChatGPT : **أثناء مرحلة التدريب المسبق، أتعلم على ملايين الجمل من الإنترنت، الكتب، المقالات...**

في كل مرة أُعطى فيها جملة ناقصة، يُطلَب مني تخمين الكلمة التالية، ويتم تصحيحي.

مثال:

"ذهب الطفل إلى"_____

أحاول أن أقول: "المدرسة" أو "النوم" أو "السوق"، بحسب ما تعلّمته.

إذا كانت الإجابة صحيحة، يُكافأ النموذج،

وإذا كانت خطأ، يُعاقب ويُصحّح داخليًا عبر "الانحدار العكسي".

👨‍🎓 الطالب: هل هذا يشبه التفكير البشري؟

🤖 ChatGPT : نوعًا ما. تشبه عملية القراءة لدى الإنسان:

أولًا ترى الكلمة، ثم تفهم معناها،

ثم تربطها بما قبلها، ثم تفهم نبرة المتكلم أو غايته.

كل خطوة تشبه طبقة من المعالجة، مجرد تجميع رياضي متتابع يضيف طبقة فوق طبقة.

👨‍🎓 الطالب: هل يمكن رؤية ما يجري داخل هذه الطبقات؟

🤖 ChatGPT: نعم، عبر تقنيات مثل:

"التمثيل البصري للتنشيط (Activation Maps) "

أو تحليل الأوزان.

الباحثون يحاولون فهم ما الذي تتعلمه كل طبقة.

👨‍🎓 الطالب: كم عدد الطبقات عادة؟

🤖 ChatGPT: يختلف حسب النموذج. مثلًا:

- BERT-base: يحتوي على ١٢ طبقة.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

- GPT-٣: يحتوي على ٩٦ طبقة،

وكلما زاد عدد الطبقات زادت القدرة الحسابية، ولكن أيضًا زادت التكلفة.

👨‍🎓 الطالب: يعني كل متجه يدخل هذه الطبقات ويتحوّل داخليًّا؟

🤖 ChatGPT: نعم، بالضبط.

كل طبقة تقرأ المتجه وتقوم بما يلي:

١. تُحلل علاقته بالمتجهات الأخرى (عبر آلية الانتباه الذاتي).

٢. تُعيد تشكيله بشكلٍ جديد.

٣. ترسله إلى الطبقة التي بعدها. وهكذا...

👨‍🎓 الطالب: هل لهذه الطبقات علاقة بجودة الإجابة؟

🤖 ChatGPT : نعم، كلما كانت الطبقات مدرّبة تدريبًا جيدًا، وتعمل بتناغم، زادت جودة الفهم، وبالتالي زادت دقة الإجابة.

👨‍🎓 الطالب: يعني كل متجه يمرّ بطبقة... ثم يُعاد تشكيله؟

🤖 ChatGPT: أحسنت!

كل طبقة تُضيف فهمًا أعمق للمتجه،

فكأن كل طبقة تكتشف بُعدًا جديدًا من معنى الكلمة.

🔍 مثال عملي:

لو كتبت لي:

"المدرسة جميلة لكن مزدحمة"

كل كلمة تتحوّل إلى متجه.

ثم:

١. الطبقة الأولى تدرك الترتيب: "مدرسة ← جميلة ← مزدحمة".

٢. الطبقة الثانية تعرف أن "جميلة" تصف "المدرسة"، و"لكن" تعارض ما بعدها.

٣. الطبقة الثالثة تفهم أن الجملة فيها حكم إيجابي + ملاحظة سلبية.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

٤. الطبقة الرابعة تُقدّر أن السائل قد يريد رأيًا أو اقتراحًا.

٥. الطبقة الأخيرة تولّد ردًا مناسبًا مثل:

"ربما يمكنك الحضور في وقت أقل ازدحامًا."

🧪 مثال:

إذا قلتَ لي: "اجتهد خالد فنجح في الامتحان"،

الطبقات الأولى تتعامل مع البنية،

والطبقات الأعلى تفهم أن هناك علاقة سببية.

📊 ملخص المفاهيم:

المفهوم	التوضيح
الطبقة	مستوى من المعالجة داخل النموذج
التراكم التمثيلي	كل طبقة تضيف تمثيلًا أعمق للكلمات
الوظائف المتنوعة	طبقات تركز على النحو، وأخرى على المعنى، وأخرى على السياق

📌 خاتمة الفصل:

الطبقات الخفية تشبه طبقات التأويل: تبدأ من السطح، وتنتهي إلى المعنى الدقيق. ولفهم كيف تنتقل الكلمة بين هذه الطبقات، دعنا نرافقها في رحلتها داخل النموذج... في الفصل القادم. هي الإطار العام الذي ينسّق كل ما سبق: من التوكنات، إلى الانتباه، إلى التنبؤ. إنها العمود الفقري للنماذج اللغوية الحديثة، وهي ما سيمهد لنا لفهم الطبقات الخفية القادمة داخل النموذج.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل الثامن: رحلة المتجه داخل النموذج

📌 توطئة الفصل:

لقد تعرّفنا على المتجه، والانتباه، والطبقات. لكن كيف تمرّ الكلمة فعليًا داخل النموذج؟ ما هي المحطات التي تعبرها؟ هذا الفصل يحاكي الرحلة الكاملة التي تبدأ من إدخال الكلمة، وتنتهي بإنتاج الجواب.

👨‍🎓 الطالب: ما أول خطوة تمر بها الكلمة داخل النموذج؟

🤖 ChatGPT: أولًا، تُحوَّل الكلمة إلى توكن(Token) ، أي تمثيل عددي مخصص. ثم يُحوَّل هذا التوكن إلى متجه باستخدام جدول مضمن.(Embedding Table)

👨‍🎓 الطالب: وماذا يحدث بعد تحويلها إلى متجه؟

🤖 ChatGPT: يُرسل المتجه إلى الطبقة الأولى من طبقات النموذج. وهناك، تبدأ رحلة المعالجة: انتباه ذاتي، ثم تغذية أمامية، ثم إرسال الناتج إلى الطبقة التالية.

👨‍🎓 الطالب: هل تتغير قيم المتجه في كل طبقة؟

🤖 ChatGPT: نعم. كل طبقة تضيف طبقة جديدة من التمثيل، بناءً على ما تعلّمته. فالمتجه يُعاد تشكيله باستمرار حتى يصل إلى تمثيل نهائي غني.

👨‍🎓 الطالب: وكيف يتم اختيار الكلمة الناتجة؟

🤖 ChatGPT: بعد أن تنتهي المعالجة في آخر طبقة، يُحوّل المتجه النهائي إلى قائمة احتمالات باستخدام دالةSoftmax ، وتُختار الكلمة صاحبة أعلى احتمال.

👨‍🎓 الطالب: إذًا الكلمة تمر بسلسلة كاملة من التحولات؟

🤖 ChatGPT: بالضبط. الكلمة تدخل كنص، ثم تتحول إلى توكن، ثم إلى متجه، ثم تمر عبر طبقات متعددة، ثم تتحول إلى احتمال، ثم تُنتج ككلمة جديدة.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

📊 ملخص المفاهيم:

المرحلة	التوضيح
التوكن	تمثيل رقمي مبدئي للكلمة
المتجه الأولي	متجه من جدول التضمين يمثل الكلمة رقمياً
التمثيل المتراكم	تعديلات متتالية على المتجه في كل طبقة
الاحتمالات النهائية	استخدام Softmax لاختيار الكلمة الأنسب

📌 خاتمة الفصل:

رحلة المتجه داخل النموذج هي جوهر التوليد اللغوي. إنها العملية التي تنقل الكلمة من مجرّد حروف إلى تمثيل ذكي، ومن تمثيل إلى قرار، ومن قرار إلى جملة. وفي الفصل القادم، سنتعرّف على كيف يتعلّم النموذج من الخطأ، عبر الانحدار العكسي. طبقات التأويل: تبدأ من السطح، وتنتهي إلى المعنى الدقيق. ولفهم كيف تنتقل الكلمة بين هذه الطبقات، دعنا نرافقها في رحلتها داخل النموذج... في الفصل القادم. هي الإطار العام الذي ينسّق كل ما سبق: من التوكنات، إلى الانتباه، إلى التنبؤ. إنها العمود الفقري للنماذج اللغوية الحديثة، وهي ما سيمهد لنا لفهم الطبقات الخفية القادمة داخل النموذج.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل التاسع: الانحدار العكسي: كيف يتعلّم النموذج من الخطأ؟

📌 توطئة الفصل:

لكي يتحسّن النموذج ويصبح أكثر دقة، يجب أن يتعلّم من أخطائه. هذا لا يحدث كما يحدث مع الإنسان بالتجربة والشعور، بل عبر عملية رياضية دقيقة تُعرف باسم "الانحدار العكسي". في هذا الفصل، نُشرّح كيف يتم تصحيح الخطأ داخل أعماق النموذج.

👨‍🎓 الطالب: ما المقصود بـ"الانحدار العكسي"؟

🤖 ChatGPT: الانحدار العكسي (Backpropagation) هو خوارزمية تُستخدم لتحديث أوزان النموذج بعد كل تجربة تعليمية، من خلال حساب الخطأ الناتج عن التنبؤ، وتوزيعه عكسيًّا على الطبقات لتعديلها.

👨‍🎓 الطالب: كيف يعرف النموذج أنه أخطأ؟

🤖 ChatGPT: أثناء التدريب، يُقارن الناتج الذي يولده النموذج بالإجابة الصحيحة (Label). ويُحسب الفرق بينهما عبر دالة خسارة (Loss Function). هذا الفرق يُمثّل مقدار الخطأ.

👨‍🎓 الطالب: وما الذي يحدث بعد حساب الخطأ؟

🤖 ChatGPT: يتم نشر هذا الخطأ إلى الوراء عبر جميع الطبقات(Backpropagation) ، وتُحسب اشتقاقات رياضية تُبيّن كيف يجب تعديل كل وزن لجعل النموذج أقل خطأً في المرة القادمة.

👨‍🎓 الطالب: ما نوع الخوارزميات المستخدمة؟

🤖 ChatGPT: أكثرها شيوعًا هو خوارزمية التدرّج العكسي مع التحسين عبر Gradient Descent. تُستخدم لتقليل قيمة الخطأ تدريجيًا.

👨‍🎓 الطالب: هل هذا يشبه التعلّم البشري؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

ChatGPT: 🤖 من حيث النتيجة، نعم: النموذج يتحسن مع الوقت. لكن من حيث الآلية، لا: لا توجد نية، ولا وعي، بل فقط تعديل رياضي تراكمي مبني على الحسابات.

📊 ملخص المفاهيم:

المفهوم	التوضيح
الانحدار العكسي	آلية لتعديل أوزان النموذج بناءً على الخطأ
دالة الخسارة(Loss)	فرق بين الناتج المتوقع والناتج الفعلي
التدرّج التنازلي	خوارزمية لتقليل الخطأ تدريجيًا

📌 خاتمة الفصل:

من دون انحدار عكسي، لا يمكن للنموذج أن يتعلم. إنه العقل الرياضي الذي يُصحح ذاته باستمرار. وفي الفصل الأخير، نضع كل هذه المفاهيم في سياقها التطوري، لنقارن بين النماذج المختلفة التي صنعت الثورة الذكية.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

الفصل العاشر: أنواع النماذج

📌 توطئة الفصل:

ما الذي ميّز GPT عن غيره؟ وكيف تطوّرت النماذج من مجرد تمثيل بسيط للكلمات إلى أنظمة توليد ذكية؟ في هذا الفصل نختم رحلتنا بمقارنة شاملة بين النماذج اللغوية الأهمWord٢Vec :، BERT، و.GPT

👨‍🎓 الطالب: ما هوWord٢Vec ؟

ChatGPT: 🤖 هو نموذج بسيط نسبيًّا تم تطويره في Google عام ٢٠١٣. يعتمد على فكرة تمثيل كل كلمة بمتجه ثابت، بناءً على سياقاتها المجاورة في النصوص. لا يأخذ السياق الكامل بعين الاعتبار، لكنه كان بداية ثورة التمثيل العددي للكلمات.

👨‍🎓 الطالب: وماذا عنBERT ؟

ChatGPT: BERT (Bidirectional Encoder Representations from Transformers) طُوّر في Google أيضًا عام ٢٠١٨. يتميّز بكونه ثنائي الاتجاه، يقرأ الجملة من الجانبين، ويولّد تمثيلًا سياقيًا عميقًا لكل كلمة.

👨‍🎓 الطالب: وهل يستخدم نفس البنية مثلGPT ؟

ChatGPT: 🤖 ليس تمامًا BERT .يستخدم المحوّل بطريقة ترميز فقط(Encoder-only) ، بينما GPT يستخدم الترميز والتوليد(Decoder-style) ، ولهذا يُستخدم GPT بشكل أفضل في المهام الإبداعية مثل الكتابة.

👨‍🎓 الطالب: ما الفرق الجوهري بين BERT وGPT؟

ChatGPT: BERT 🤖 يفهم، و GPTيُنتج. الأول يُستخدم لتحليل النصوص، والثاني لتوليدها . BERTيتوقف عند التفسير، أما GPT فيبني استجابات جديدة.

👨‍🎓 الطالب: هل ٤-GPT مثل النماذج السابقة؟

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

ChatGPT: 🤖 لا ٤-GPT. يُعد من النماذج متعددة الوسائط(Multimodal) ، وقادر على التعامل مع نصوص وصور وصوت. كما يتميّز بذاكرة أفضل وتوجيه سلوكي أدق.

📊 ملخص المفاهيم:

النموذج	المزايا الرئيسية	الاستخدامات
Word٢Vec	تمثيل ثابت للكلمات بناءً على السياق	تحليل بسيط للكلمات
BERT	تمثيل سياقي ثنائي الاتجاه	التصنيف، الاستخراج، الفهم العميق
GPT	توليد إبداعي أحادي الاتجاه، متعدد الأدوار	الكتابة، المحادثة، الترجمة

📌 خاتمة الفصل:

كل نموذج هو خطوة في مسار التطور نحو فهم اللغة. بدأنا بالتمثيل الثابت، ثم أضفنا السياق، ثم انتقلنا إلى التوليد الذكي. لكن هذه ليست نهاية الطريق، فكل يوم يظهر جديد. والفهم الحقيقي لما يجري داخل هذه النماذج هو ما يمكّننا من الاستفادة الواعية، لا الانبهار الساذج.، لا يمكن للنموذج أن يتعلم. إنه العقل الرياضي الذي يُصحح ذاته باستمرار. وفي الفصل الأخير، نضع كل هذه المفاهيم في سياقها التطوري، لنقارن بين النماذج المختلفة التي صنعت الثورة الذكية.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

خاتمة الكتاب:

ها قد انتهت رحلتنا من الكلمة إلى المتجه، ومن المتجه إلى السياق، ومن السياق إلى التوليد الذكي. عبر عشرة فصول مترابطة، كشفنا عن البنية الخفية التي تجعل النماذج اللغوية قادرة على قراءة اللغة، ومعالجتها، وإنتاجها، بطريقة تبدو لنا أقرب ما تكون إلى الفهم.

لكننا أيضًا أعدنا التأكيد أن ما يجري ليس فكرًا واعيًا، ولا إدراكًا إنسانيًا، بل محاكاة رياضية مركّبة، تتقن الحساب، وتُحسن التقليد، لكنها لا تزال بلا وعي ولا قصد ولا شعور.

هذا الكتاب لا يكتفي بعرض تقنيّ، بل يدعو إلى التأمل النقديّ، والتساؤل المستمر: هل ما نراه من "ذكاء اصطناعي" هو بالفعل ذكاء؟ أم هو مجرّد صدى لتحسينات إحصائية متقدمة؟

وبينما تتطوّر النماذج، وتزداد تعقيدًا وذكاءً ظاهريًا، تبقى مسؤوليتنا أن نفهمها بعمق، وأن نوجّهها استخدامًا وأخلاقًا، وأن نستبصر آثارها على المعرفة، والتعليم، والهوية، والمستقبل.

فالفهم هو السلاح الأول في زمن الذكاء الاصطناعي، وهذا الكتاب خطوة متواضعة في طريق هذا الفهم.

قائمة المراجع

مراجع الفصل الأول: التمثيل المتّجه

١. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (٢٠١٣). *Efficient Estimation of Word Representations in Vector Space*. arXiv preprint arXiv:١٣٠١.٣٧٨١.

٢. Pennington, J., Socher, R., & Manning, C. D. (٢٠١٤). *GloVe: Global Vectors for Word Representation*. EMNLP.

٣. Goldberg, Y. (٢٠١٦). *A Primer on Neural Network Models for Natural Language Processing*. Journal of Artificial Intelligence Research.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

٤. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing* (٣rd ed.). Pearson. – الفصل المتعلق بـ Word Embeddings.

٥. Vaswani, A. et al. (٢٠١٧). *Attention is All You Need*. NeurIPS. –لتوضيح الخلفية التي مهّدت للتضمينات السياقية في الفصول اللاحقة.

مراجع الفصل الثاني: التوكنات والتكلفة

١. Sennrich, R., Haddow, B., & Birch, A. (٢٠١٦). *Neural Machine Translation of Rare Words with Subword Units*. ACL.

٢. Radford, A. et al. (٢٠١٩). *Language Models are Unsupervised Multitask Learners*. OpenAI Technical Report.

٣. Brown, T. et al. (٢٠٢٠). *Language Models are Few-Shot Learners*. arXiv preprint arXiv:٢٠٠٥.١٤١٦٥.

٤. OpenAI Documentation. (٢٠٢٣). *Tokenization and Pricing Overview*. Retrieved from: https://platform.openai.com

٥. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing* (٣rd ed.). – الفصل المتعلق بـ Tokenization and Cost.

مراجع الفصل الثالث: التضمين السياقي

١. Peters, M. et al. (٢٠١٨). *Deep contextualized word representations*. NAACL-HLT.

٢. Devlin, J. et al. (٢٠١٩). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv preprint arXiv:١٨١٠.٠٤٨٠٥.

٣. Liu, Y. et al. (٢٠١٩). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. arXiv preprint arXiv:١٩٠٧.١١٦٩٢.

٤. Tenney, I. et al. (٢٠١٩). *What do you learn from context? Probing for sentence structure in contextualized word representations*. ICLR.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

٥. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing* (٣rd ed.). – الفصل المتعلق بـ Contextual Embeddings.

مراجع الفصل الرابع: التنبؤ الإحصائي

١. Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (٢٠٠٣). *A Neural Probabilistic Language Model*. Journal of Machine Learning Research.

٢. Radford, A. et al. (٢٠١٩). *Language Models are Unsupervised Multitask Learners*. OpenAI.

٣. Brown, T. et al. (٢٠٢٠). *Language Models are Few-Shot Learners*. arXiv:٢٠٠٥.١٤١٦٥.

٤. Golding, A. R., & Roth, D. (١٩٩٩). *A Winnow-Based Approach to Context-Sensitive Spelling Correction*. Machine Learning Journal.

٥. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing*. ٣rd ed. – الفصول المتعلقة بالتنبؤ الاحتمالي و Softmax.

مراجع الفصل الخامس: الانتباه الذاتي

١. Vaswani, A. et al. (٢٠١٧). *Attention is All You Need*. Advances in Neural Information Processing Systems (NeurIPS).

٢. Clark, K., Khandelwal, U., Levy, O., & Manning, C. D. (٢٠١٩). *What Does BERT Look at? An Analysis of BERT's Attention*. ACL.

٣. Bahdanau, D., Cho, K., & Bengio, Y. (٢٠١٥). *Neural Machine Translation by Jointly Learning to Align and Translate*. ICLR.

٤. Rogers, A., Kovaleva, O., & Rumshisky, A. (٢٠٢٠). *A Primer in BERTology: What We Know About How BERT Works*. TACL.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

٥. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing* (٣rd ed.). – الفصل المتعلق بـ Self-Attention and Transformers.

مراجع الفصل الرابع: التنبؤ الإحصائي

١. Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (٢٠٠٣). *A Neural Probabilistic Language Model*. Journal of Machine Learning Research.

٢. Radford, A. et al. (٢٠١٩). *Language Models are Unsupervised Multitask Learners*. OpenAI.

٣. Brown, T. et al. (٢٠٢٠). *Language Models are Few-Shot Learners*. arXiv:٢٠٠٥.١٤١٦٥.

٤. Golding, A. R., & Roth, D. (١٩٩٩). *A Winnow-Based Approach to Context-Sensitive Spelling Correction*. Machine Learning Journal.

٥. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing*. ٣rd ed. – الفصول المتعلقة بالتنبؤ الاحتمالي و Softmax.

مراجع الفصل السادس: معمارية المحوّل

١. Vaswani, A. et al. (٢٠١٧). *Attention is All You Need*. NeurIPS.

٢. Alammar, J. (٢٠١٨). *The Illustrated Transformer*. Blog post: https://jalammar.github.io/illustrated-transformer/

٣. Wolf, T. et al. (٢٠٢٠). *Transformers: State-of-the-Art Natural Language Processing*. EMNLP.

٤. Raffel, C. et al. (٢٠٢٠). *Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer*. JMLR.

٥. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing* (٣rd ed.). – فصل Transformers and Encoder-Decoder Architectures.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

مراجع الفصل السابع: الطبقات الخفية

١. Raghu, M. et al. (٢٠١٧). *SVCCA: Singular Vector Canonical Correlation Analysis for Deep Learning Dynamics and Interpretability*. NeurIPS.

٢. Alain, G., & Bengio, Y. (٢٠١٦). *Understanding Intermediate Layers Using Linear Classifier Probes*. ICLR.

٣. Tenney, I., Das, D., & Pavlick, E. (٢٠١٩). *BERT Rediscovers the Classical NLP Pipeline*. ACL.

٤. Rogers, A., Kovaleva, O., & Rumshisky, A. (٢٠٢٠). *A Primer in BERTology: What We Know About How BERT Works*. TACL.

٥. Olah, C., Satyanarayan, A., et al. (٢٠١٨). *The Building Blocks of Interpretability*. Distill.pub.

مراجع الفصل الثامن: رحلة المتجه داخل النموذج

١. Vaswani, A. et al. (٢٠١٧). *Attention is All You Need*. NeurIPS.

٢. Alammar, J. (٢٠١٨). *The Illustrated Transformer*. jalammar.github.io

٣. Olah, C. et al. (٢٠١٨). *The Building Blocks of Interpretability*. Distill.

٤. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing*. ٣rd ed.

٥. Wolf, T. et al. (٢٠٢٠). *Transformers: State-of-the-Art NLP*. EMNLP.

مراجع الفصل التاسع: الانحدار العكسي

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)

١. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (١٩٨٦). *Learning representations by back-propagating errors*. Nature.
٢. Goodfellow, I., Bengio, Y., & Courville, A. (٢٠١٦). *Deep Learning*. MIT Press.
٣. LeCun, Y., Bottou, L., Orr, G., & Müller, K. (١٩٩٨). *Efficient BackProp*. Neural Networks.
٤. Jurafsky, D., & Martin, J. H. (٢٠٢٣). *Speech and Language Processing*. ٣rd ed.
٥. Nielsen, M. (٢٠١٥). *Neural Networks and Deep Learning*. Online Book.

📚 **مراجع الفصل العاشر: أنواع النماذج**

١. Mikolov, T. et al. (٢٠١٣). *Efficient Estimation of Word Representations in Vector Space*. arXiv.
٢. Devlin, J. et al. (٢٠١٩). *BERT: Pre-training of Deep Bidirectional Transformers*. arXiv.
٣. Radford, A. et al. (٢٠١٨–٢٠٢٣). *GPT Model Reports*. OpenAI.
٤. Raffel, C. et al. (٢٠٢٠). *T٥: Exploring the Limits of Transfer Learning*. JMLR.
٥. Liu, Y. et al. (٢٠١٩). *RoBERTa: A Robustly Optimized BERT*. arXiv.

يُمنع نسخ هذا المحتوى أو أي جزء منه دون إذن خطي من المؤلف. لمزيد من المعلومات (النقر على الرابط)