

استخدام الذكاء الاصطناعي لخدمة اللغة العربية: تطبيقات لتعلم الآلة العميق

الدكتور فيث علي عنبدة
أستاذ هندسة الحاسوب في الجامعة الأردنية
مدير مشروع تطوير مناهج الذكاء الاصطناعي والروبوتات
الثلاثاء 2 تشرين الثاني 2021
اتحاد الجامعات العربية

1

الأجندة

- مقدمة عن الذكاء الاصطناعي وتعلم الآلة
- إنجازات ومحددات الذكاء الاصطناعي المعاصر
- معالجة اللغات الطبيعية
- تطبيقات في معالجة اللغة العربية

2

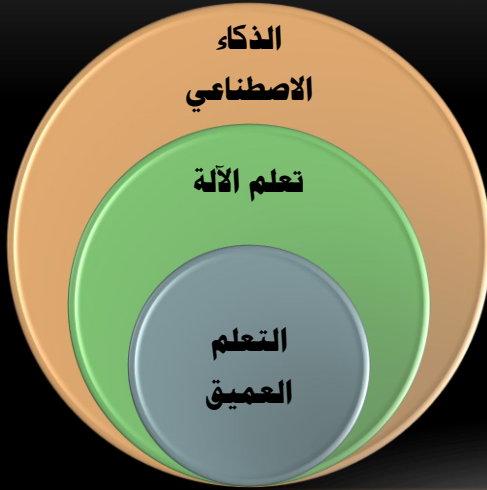
مقدمة عن الذكاء الاصطناعي

- **الذكاء (Intelligence)** هو القدرة على إنجاز أهداف معقدة
- **الذكاء الاصطناعي (Artificial Intelligence)** هو الذكاء غير البيولوجي المصنع أو المبرمج من الإنسان
- **الذكاء الضيق (Narrow Intelligence)** هو القدرة على إنجاز مجموعة محددة وضيقة من الأهداف، مثل الفوز في لعبة شطرنج أو قيادة سيارة بأمان

مقدمة عن الذكاء الاصطناعي

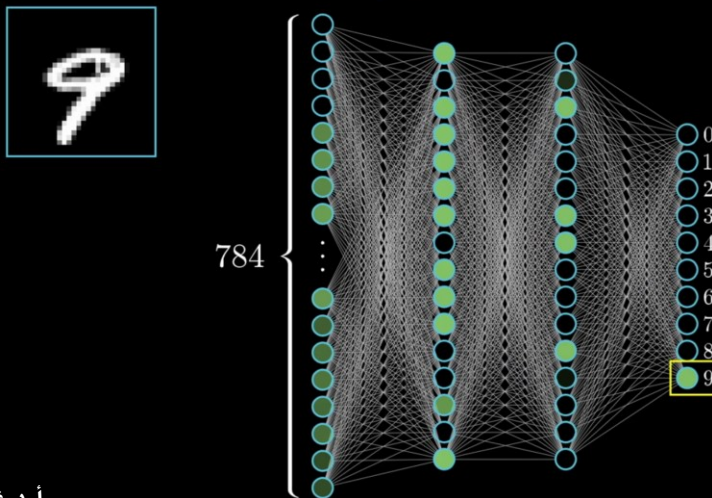
- **الذكاء العام (General Intelligence)** مثل الموجود عند الإنسان هو القدرة على إنجاز أي هدف بما في ذلك القدرة على التعلم
- يسعى كثير من المختصين والشركات في الوقت الحاضر إلى تطوير ذكاء اصطناعي عام
- **الذكاء الخارق (Super AI)** سيتبع الذكاء الاصطناعي العام وهو ذكاء عام فوق مستوى الذكاء البشري

مقدمة عن تعلم الآلة



- **تعلم الآلة (Machine Learning)** هو قسم من طرق الذكاء الاصطناعي باستعمال خوارزميات يتحسن أداؤها (تتعلم) من تدريبها على بيانات
- **والتعلم العميق (Deep Learning)** هو قسم من طرق تعلم الآلة يعتمد على العديد من طبقات الشبكات العصبية الاصطناعية

الشبكات العصبونية العميقة



مقدمة عن تعلم الآلة

• أنواع تعلم الآلة الرئيسية

- | | |
|---------------|------------------|
| Supervised | 1. تعلم موجه |
| Unsupervised | 2. تعلم غير موجه |
| Reinforcement | 3. تعلم معزز |

الأجندة

- مقدمة عن الذكاء الاصطناعي وتعلم الآلة
- إنجازات ومحددات الذكاء الاصطناعي المعاصر
- معالجة اللغات الطبيعية
- تطبيقات في معالجة اللغة العربية

لماذا ينجح الذكاء الاصطناعي الآن؟

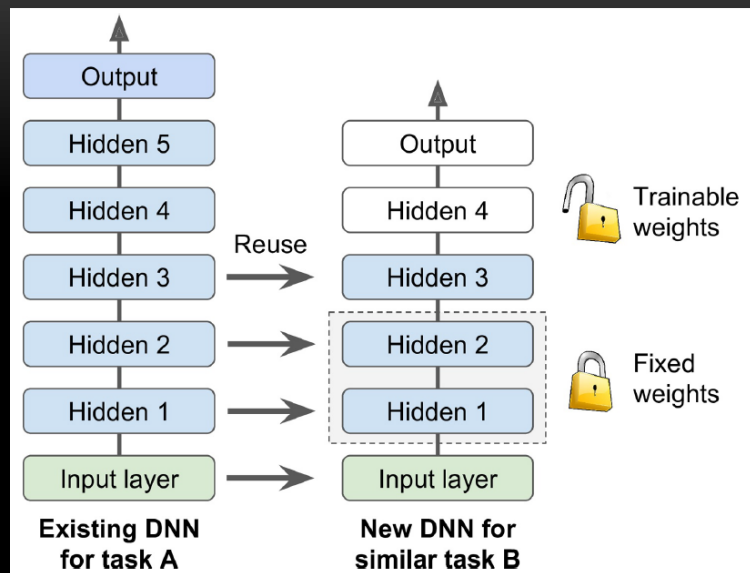
- توفر البيانات اللازمة لتعلم الآلة
- تطوير خوارزميات
- زيادة سرعة الحواسيب

إنجازات الذكاء الاصطناعي المعاصر

1. التعرف على صورنا وأصواتنا
2. تقديم توصيات من صاحب وماذا نشاهد ونقرأ
3. مساعدتنا في البحث واسترجاع المعلومات
4. الترجمة ومعالجة اللغات الطبيعية
5. تسيير المركبات المستقلة
6. تطبيقات الأمان واكتشاف الخروقات والتجاوزات وشبكات الأشياء

إنجازات الذكاء الاصطناعي المعاصر

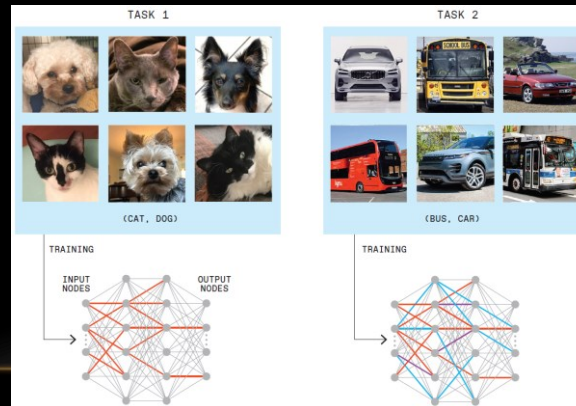
7. الذكاء الاصطناعي يعطي حلول رخيصة ذات دقة مقبولة
8. توفر نماذج مدربة جاهزة يمكن للمبرمج تحميلها واستخدامها
9. **نقل التعلم (Transfer Learning)** يمكن إعادة تدريب النماذج الجاهزة لحل مسائل جديدة قريبة من المسألة التي يظها النموذج الأصلي



نقل التعلم

محددات الذكاء الاصطناعي المعاصر

1. الذكاء الاصطناعي الموجود حالياً هو ذكاء ضيق



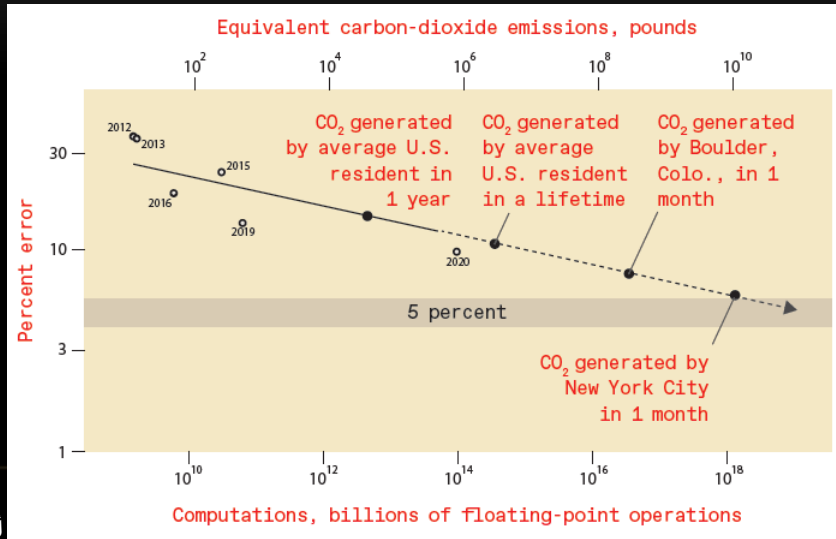
محددات الذكاء الاصطناعي المعاصر

1. الذكاء الاصطناعي الموجود حالياً هو ذكاء ضيق

2. تعلم الآلة يحتاج قواعد بيانات كبيرة

3. تعلم الآلة يأخذ أوقات تعلم طويلة

تعلم الآلة يأخذ أوقات تعلم طويلة



أ. د. غيث عنبدة

15

15

محددات الذكاء الاصطناعي المعاصر

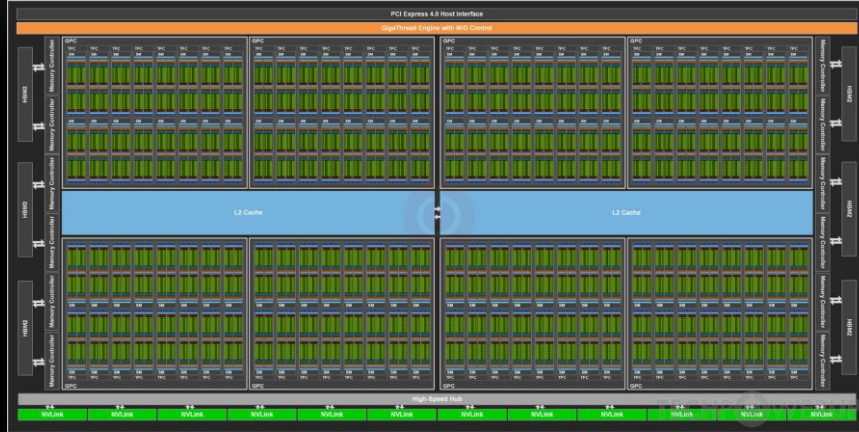
1. الذكاء الاصطناعي الموجود حالياً هو ذكاء ضيق
2. تعلم الآلة يحتاج قواعد بيانات كبيرة
3. تعلم الآلة يأخذ أوقات تعلم طويلة
4. تعلم الآلة يحتاج حواسيب خاصة

أ. د. غيث عنبدة

16

16

**Nvidia GA100 GPU: 826 mm² chip, 54 billion transistors, 108 SM,
6,912 FP32 CUDA cores, 40 GB memory**



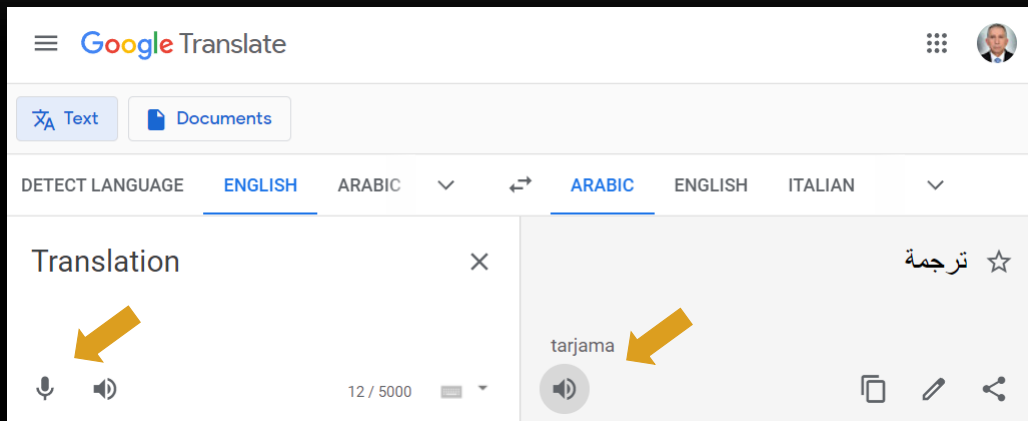
الأجندة

- مقدمة عن الذكاء الاصطناعي وتعلم الآلة
- إنجازات ومحددات الذكاء الاصطناعي المعاصر
- معالجة اللغات الطبيعية
- تطبيقات في معالجة اللغة العربية

معالجة اللغات الطبيعية

- **معالجة اللغات الطبيعية (Natural Language Processing)** هو فرع من الذكاء الاصطناعي، يعنى بتطوير قدرات الحواسيب لمعالجة وتحليل وفهم والتخاطب باللغات الطبيعية مثل الإنجليزية والعربية.
- اعتمدت معالجة اللغات الطبيعية سابقاً على المعاجم والبرمجة القائمة على القواعد.
- تتقدم معالجة اللغات الطبيعية الآن سريعاً باستعمال تقنيات التعلم العميق.
- يوجد الآن تطبيقات ناجحة لمعالجة اللغات تعتمد على التعلم العميق.

الترجمة والتعرف على الكلام وتوليف الكلام



المساعد الإلكتروني الافتراضي



أ. د. غيث عبيدة

21

21

نظام حاسوبي للإجابة على الأسئلة



أ. د. غيث عبيدة

22

22

تطبيقات أخرى للتعلم العميق

- كشف البريد العشوائي
- تحليل المشاعر وكشف خطاب الكراهية
- التعرف على الكيانات المسماة والتحليل الصرفي والنحوي
- تصنيف وتقييم النصوص
- تلخيص النصوص وتأليف القصص الإخبارية

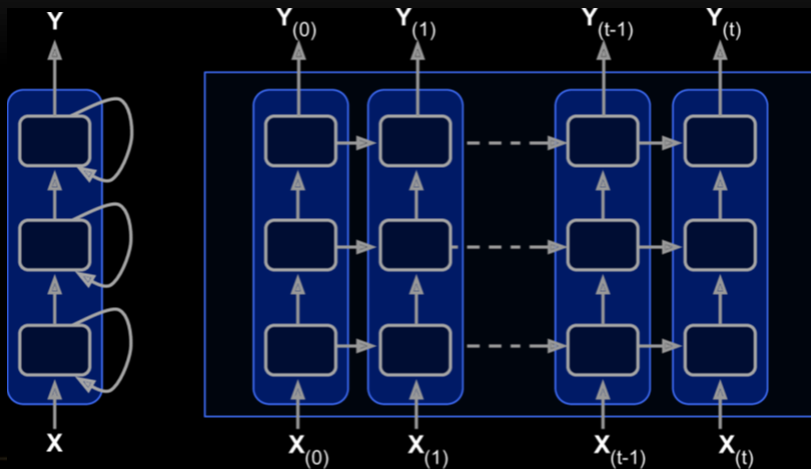
الأجندة

- مقدمة عن الذكاء الاصطناعي وتعلم الآلة
- إنجازات ومحددات الذكاء الاصطناعي المعاصر
- معالجة اللغات الطبيعية
- تطبيقات في معالجة اللغة العربية

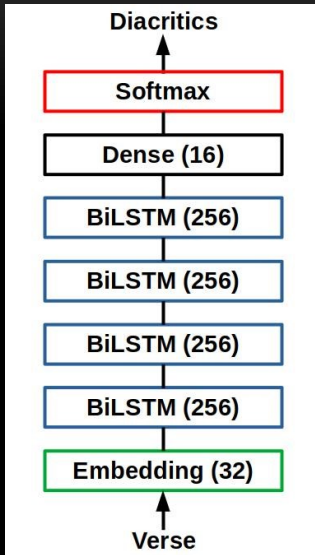
اهتماماتنا البحثية في معالجة اللغة العربية

- التعرف الآلي على النصوص العربية المطبوعة والمكتوبة
- روبوتات الدردشة
- تشكيل النصوص العربية
- تصحيح الأخطاء الإملائية والترجمة إلى لغة سليمة
- تصنيف وتشكيل الشعر العربي

الشبكات العصبية المتكررة العميقة



تشكيل النصوص العربية



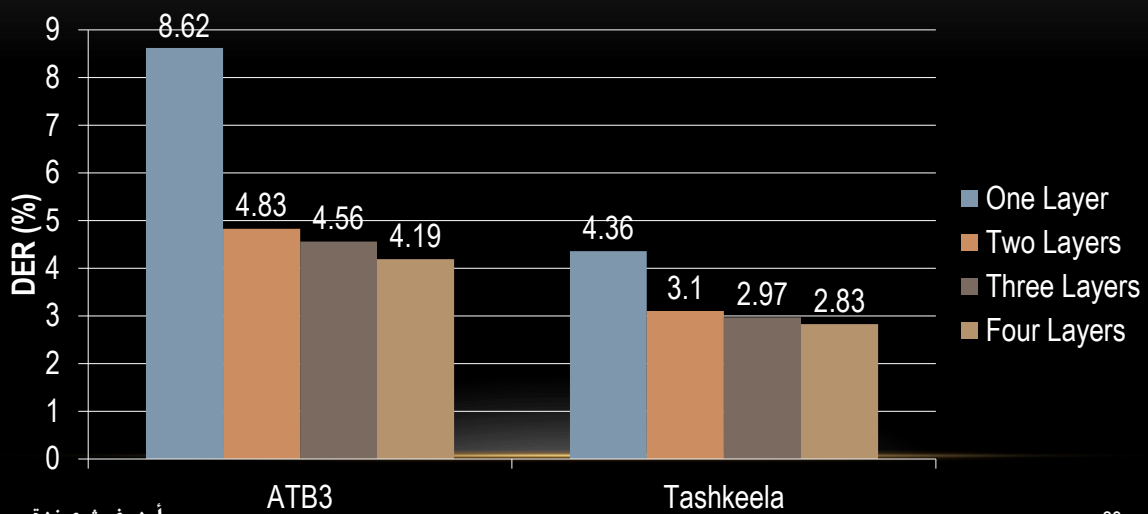
كتب الطالب رسالة
كَتَبَ الطَّالِبُ رِسَالَةً

أ. د. غيث عبيدة

27

27

الشبكة الأعمق أدق

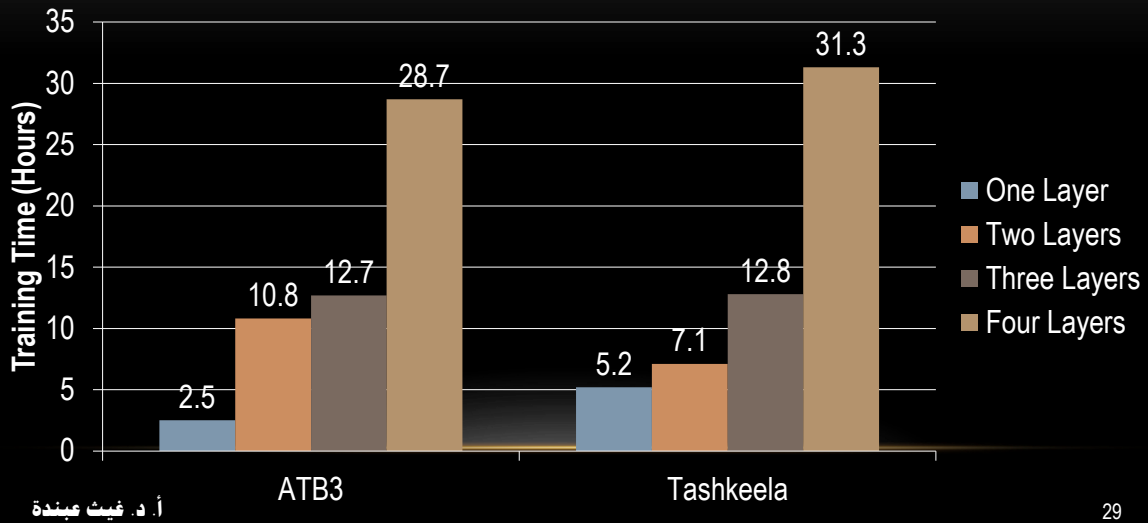


أ. د. غيث عبيدة

28

28

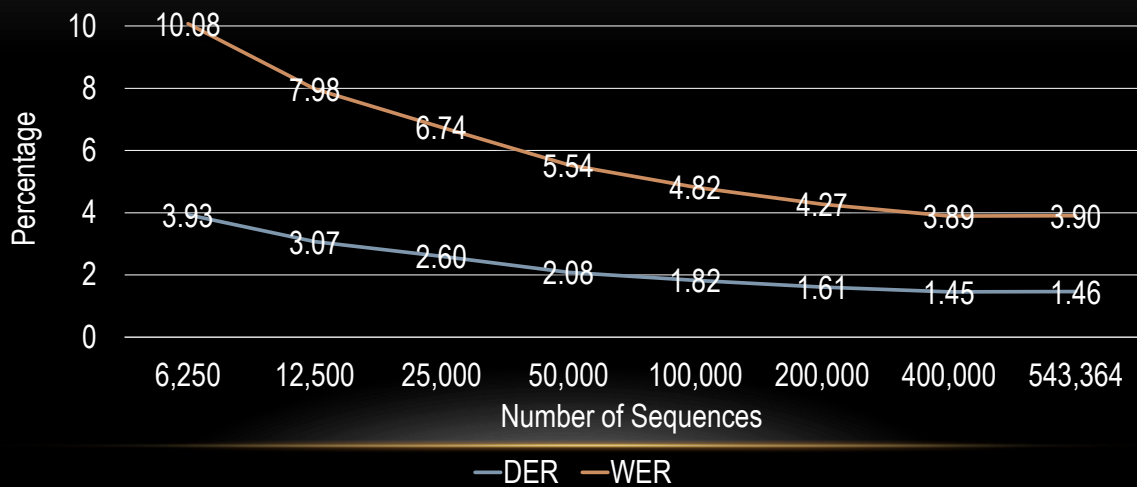
الشبكة الأعمق تحتاج زمن تدريب أطول



29

29

الدقة تتحسن عند التدريب على أمثلة أكثر



30

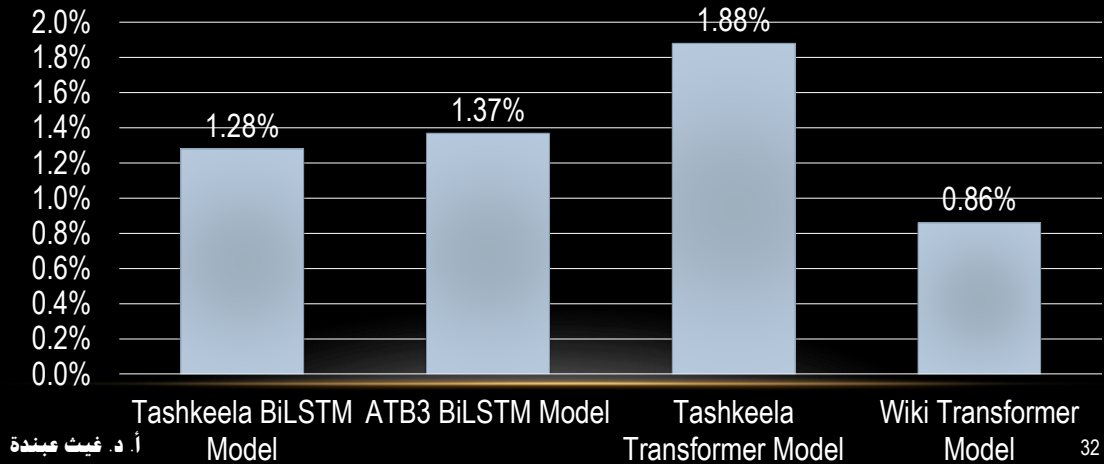
30

تصحيح الأخطاء الإملائية الشائعة

الصواب	الخطأ
جاء أبناؤهم	جاء ابنائهم
خطأ إملائي	خطأ أملائي
رجعوا كلهم صباحاً	رجعوا كلهم صباحاً
رمى العصا	رما العصى
مرت على المدرسة	مرة على المدرسه

دقة عالية مع الشبكات الحديثة والبيانات الكبيرة

CER for Test200



الشعر العربي

الرقم	البحر	الدائرة	تفعيلات الشطر الواحد
1	طويل	1	فَعولن مفاعيلن فَعولن مفاعيلن
2	كامل	2	مِفاعِلن مِفاعِلن مِفاعِلن
3	بسيط	1	مِستَفعلِن فاعِلِن مِستَفعلِن فاعِلِن
4	خفيف	4	فاعِلاتِن مِستَفعلِن فاعِلاتِن
5	واقر	2	مفاعِلتِن مفاعِلتِن فَعولِن
6	رجز	3	مِستَفعلِن مِستَفعلِن مِستَفعلِن
7	رمل	3	فاعِلاتِن فاعِلاتِن فاعِلاتِن
8	مقارب	5	فَعولِن فَعولِن فَعولِن فَعولِن فَعولِن
9	سريع	4	مِستَفعلِن مِستَفعلِن فاعِلِن
10	منسرح	4	مِستَفعلِن مفعولات مِستَفعلِن
11	مجتث	4	مِستَفعلِن فاعِلاتِن
12	مديد	1	فاعِلاتِن فاعِلِن فاعِلاتِن
13	هزج	3	مفاعِلِن مفاعِلِن
14	مبتدأرك	5	فَعولِن فَعولِن فَعولِن فَعولِن
15	مقتضب	4	مفعولات مِستَفعلِن
16	مضارع	4	مفاعِلِن فاعِلاتِن

أ. د. غيث عبيدة

33

33

تصنيف القصائد حسب البحر

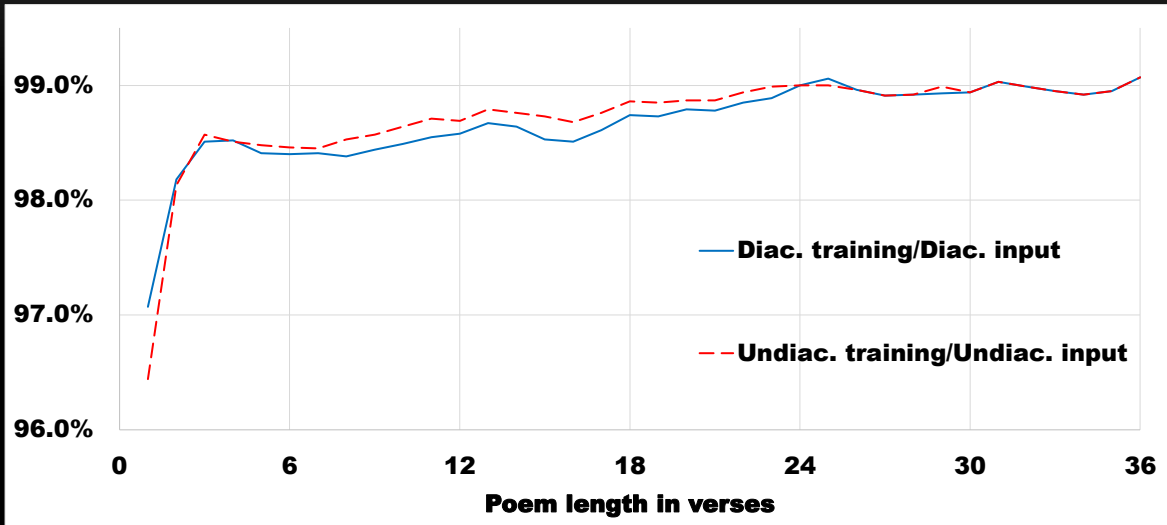


أ. د. غيث عبيدة

34

34

دقة تصنيف القصائد



أ. د. غيث عبيدة

35

35

تشكيل الشعر العربي

لعمرك ما الرزية فقد مال - ولا شاة تموت ولا بعير

ولكن الرزية فقد فذ - يموت لموته خلق كثير

لعمرك ما الرزِيَّةُ فُقِدُ مالٍ - ولا شاةٌ تَمُوتُ ولا بَعِيرُ

ولكنَّ الرزِيَّةَ فُقِدُ فذٍ - يَمُوتُ لِمَوْتِهِ خَلْقٌ كَثِيرُ

أ. د. غيث عبيدة

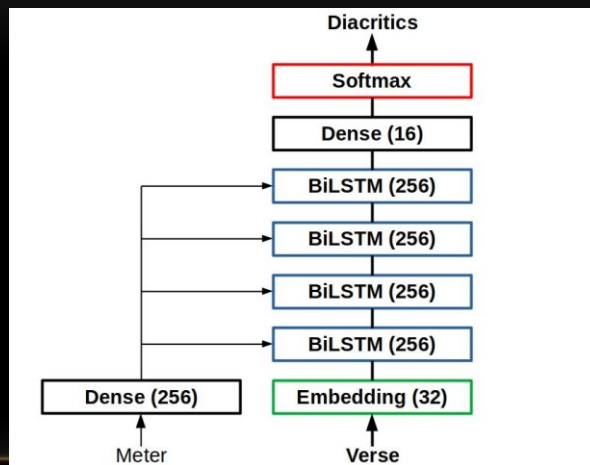
36

36

تشكيل الشعر العربي

نسبة الخطأ	عدد كلماتها	قاعدة البيانات
2.0%	2.3 مليون	نثر تراثي
2.5%	0.3 مليون	نثر حديث
6.1%	3.5 مليون	شعر

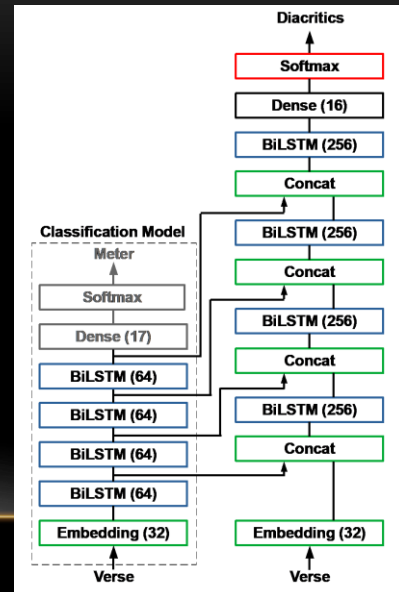
الاستفادة من معرفة البحر



الاستفادة من تفاصيل معرفة البحر ونقل التعلم

نسبة الخطأ	فترات التدريب	الشبكة المستعملة
5.2%	26	الأساسية
4.6%	26	مع معلومات البحر
3.5%	12+5+1	مع نقل التعلم

أ. د. فيث عبيدة



39

39

ملخص

- مقدمة عن الذكاء الاصطناعي وتعلم الآلة
- إنجازات ومحددات الذكاء الاصطناعي المعاصر
- معالجة اللغات الطبيعية
- تطبيقات في معالجة اللغة العربية

أ. د. فيث عبيدة

40

40

شكراً

- Email: abandah@ieee.com
- Facebook: [gheith.abandah](https://www.facebook.com/gheith.abandah)
- Twitter: [@abandah](https://twitter.com/abandah)
- LinkedIn: [gheith-abandah](https://www.linkedin.com/company/gheith-abandah)
- Website: <http://www.abandah.com/gheith>