256

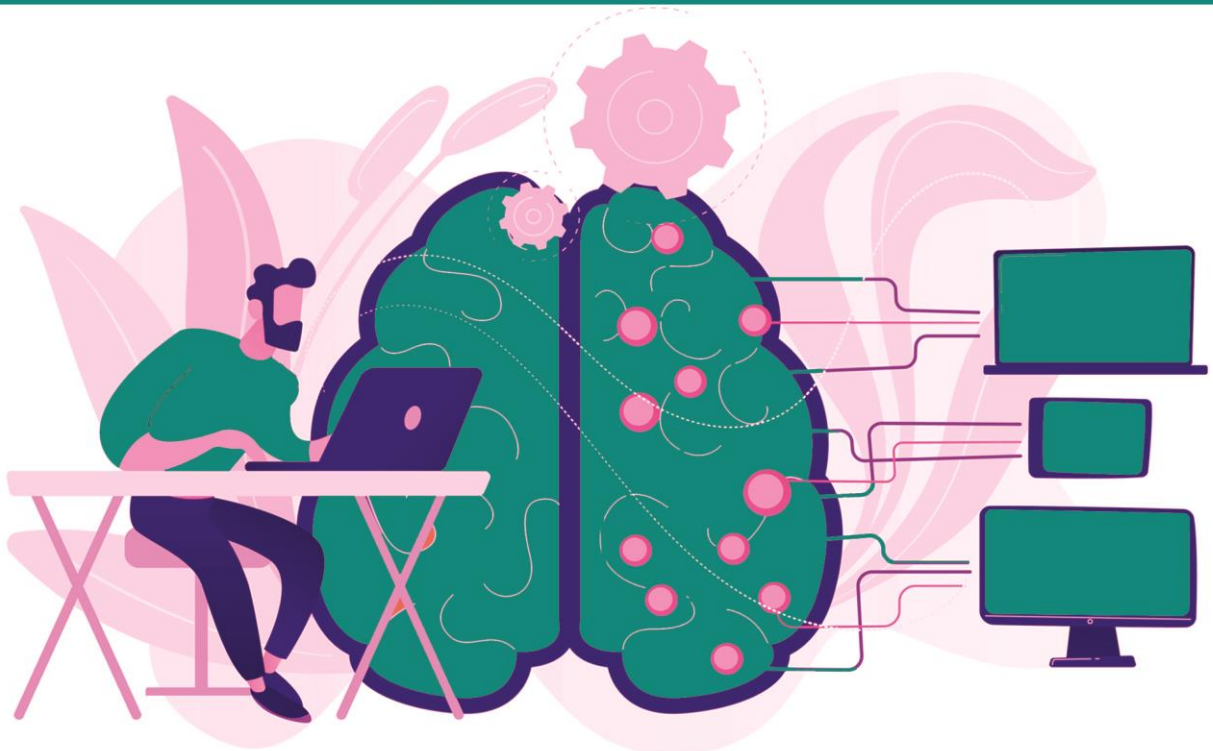
سؤال وجواب

في

التعلم الآلي والتعلم العميق وعلم البيانات

اعداد:

د. علاء طعيمة



بـــمه تعالى

256

ـــؤال وجواب في التعلم الآلي والتعلم العميق وعلم البيانات

اعداد:
د. علاء طعيمة

مقدمة المؤلف

يستهلك الباحث الكثير من الوقتفي البحث عن إجابات لأسئلة حول التعلم الآلي والتعلم العميق وعلم البيانات تندرج من السهولة إلى الأكثر صعوبة وقد تكون شديدة التنوع مما يزيد البحث عنها تعقيداً. لذا أصدرنا هذه الكتاب التي يضع بين يدي الباحث الإجابة عن الأسئلة الأكثر شيوعًا حول الجوانب المختلفة المتعلقة بالتعلم الآلي والتعلم العميق وعلم البيانات تسهيلاً للتعلم وتوفيراً للجهد.

لقد حاولت قدر المستطاع ان اجمع الأسئلة الأكثر طرحاً مع الأجوبة المناسبة والكافية، ومع هذا يبقى عملاً بشرياً يحتمل النقص، فاذا كان لديك أي ملاحظات حول هذا الكتاب، فلا تتردد بمراسلتنا عبر بريدنا الالكتروني alaa.taima@qu.edu.iq .

نأمل ان يساعد هذا الكتاب كل من يريد ان يدخلفي مجالات التعلم الآلي والتعلم العميق وعلم البيانات ومساعدة القارئ العربي على تعلم هذا المجالات. اسأل الله التوفيقفي هذا العمل لأثراء المحتوى العربي الذي يفتقر أشد الافتقار إلى محتوى جيد ورصينفي مجال التعلم الآلي والتعلم العميق وعلم البيانات. ونرجو لك الاستمتاع مع الكتاب ولا تنسونا من صالح الدعاء.

د. علاء طعيمة

كلية علوم الحاسوب وتكنولوجيا المعلومات

جامعة القادسية

العراق

المحتويات

س1/ ما هو نقل التعلم Transfer Learning في التعلم العميق؟ 14

س2/ ما هو نقل التعلم لـ CNN؟ 14

س3/ ما هو نقل التعلم مع المثال المناسب؟ 14

س4/ لماذا نحتاج نقل التعلم؟ 14

س5/ ما هو الفرق بين نقل التعلم transfer learning والضبط الدقيق fine-tuning؟ 14

س6/ ما هو الفرق بين الشبكة العصبية التلافيفية CNN و نقل التعلم Transfer Learning؟ 14

س7/ هل نقل التعلم دائمًا أفضل؟ 15

س8/ ماذا نعني بتجميد الطبقات freeze layers؟ 15

س9/ في أي الحالات يمكننا استخدام نقل التعلم؟ 15

س10/ ما هي عيوب نقل التعلم؟ 15

س11/ ما فائدة نقل التعلم؟ 15

س12/ ما هي مبادئ نقل التعلم؟ 15

س13/ ما هي الشبكة العصبية التلافيفية CNN؟ 16

س14/ هل CNN خوارزمية أم نموذج؟ 16

س15/ هل CNN خاضعة للإشراف supervised أم غير خاضعة للإشراف unsupervised؟ 16

س16/ هل تستخدم CNN للصور فقط؟ 16

س17/ هل يمكن استخدام CNN للبيانات الرقمية numerical data؟ 16

س18/ هل يمكن استخدام CNN لغير الصور non-image ؟ 16

س19/ لماذا تعد CNN أفضل من KNN؟ 16

س20/ كيف يمكنني برمجة CNN من البداية؟ 17

س21/ هل يمكننا استخدام CNN لتصنيف النص text classification؟ 17

س22/ هل يمكننا استخدام CNN لتحليل المشاعر sentiment analysis؟ 17

س23/ لماذا تعتبر CNN أفضل لتصنيف الصور؟ 17

س24/ هل CNN أفضل من SVM لتصنيف الصور؟ 17

س25/ ما هي تطبيقات CNN؟ 17

س26/ هل يمكنني استخدام CNN في السلاسل الزمنية time series؟ 18

س27/ هل رؤية الكمبيوتر هي نفسها CNN؟ 18

س28/هل Yolo مشتق من CNN؟ 18
س29/هل CNN مناسبة لـ NLP؟ 18
س30/هل يمكن استخدام CNN لاختيار الميزة feature selection؟ 18
س31/هل يمكننا استخدام CNN لاستخراج الميزة؟ 18
س32/ما هي فائدة CNN؟ 18
س33/لماذا يستخدم Max pooling في CNN؟ 19
س34/لماذا نفضل CNN على ANN بالنسبة لبيانات الصورة كمدخلات؟ 19
س35/كيف يمكنني إيقاف الضبط الزائد overfitting في CNN؟ 19
س36/هل يقلل التجميع pooling من فرط التعلم overfitting في CNN؟ 19
س37/ماذا تفعل R-CNN؟ 19
س38/هل R-CNN تعلماً عميقاً؟ 20
س39/ما هو المحسن Optimizer الأفضل لشبكة CNN؟ 20
س40/ما هو معدل التعلم learning rate في CNN؟ 20
س41/ما هي أفضل طريقة لاختيار معمارية CNN؟ 20
س42/ما هو نموذج CNN الأفضل لتصنيف الصور image classification؟ 20
س43/ما هو استخدام U-Net؟ 20
س44/هل U-Net هو نوع من CNN؟ 21
س45/هل U-Net تلافيفية بالكامل fully convolutional؟ 21
س46/ما هو اكثر نماذج التعلم العميق للتعرف على الوجوه face recognition شيوعاً؟ . 21
س47/ما هو Deepface؟ 21
س48/ما هو DeepDream؟ 21
س49/كيف يخلق الذكاء الاصطناعي الفن؟ 22
س50/ما هو Midjourney؟ 22
س51/لماذا تعتبر CNN الأفضل للتعرف على الوجوه؟ 22
~~س52/لماذا تعتبر CNN الأفضل للتعرف على الوجوه؟~~ 22
س53/ما الخوارزمية المستخدمة في التعرف على الوجوه في التعلم الآلي ML؟ 22
س54/ما هو نموذج CNN المدرب مسبقاً pre trained model الأفضل للتعرف على الوجوه؟ 22
س55/ما الخوارزمية المستخدمة للتعرف على الوجوه في بايثون؟ 23
~~س56/ما الخوارزمية المستخدمة للتعرف على الوجوه في بايثون؟~~ 23

س57/ لماذا تعتبر ResNet جيدة لتصنيف الصور؟ 23
س58/ ما هي لغة البرمجة الأفضل للتعرف على الوجه Face recognition؟ 23
س59/ ما هي الطبقة الكثيفة Dense Layer في CNN؟ 23
س60/ ما هي طبقة Softmax في CNN؟ 23
~~س61/ لماذا نقوم بتسوية البيانات data normalization في CNN؟~~ 24
س62/ لماذا يتم استخدام ReLU في الطبقات المخفية؟ 24
س63/ هل CNN هو الأفضل لاستخراج الميزات؟ 24
س64/ هل تحتاج CNN إلى اختيار الميزة feature selection؟ 24
س65/ ما هو الفرق بين الطبقة التلافيفية convolution layer وطبقة كثيفة dense layer؟ 24
س66/ ما المقصود بالفترة epoch في CNN؟ 24
س67/ هل softmax دالة تنشيط activation function ام دالة خسارة (خطأ) loss function؟ 25
س68/ ما هو حجم الدفعة batch size في CNN؟ 25
س69/ لماذا نحتاج الفترة epoch في التعلم الآلي والعميق؟ 25
س70/ هل المزيد من الفترات more epochs أفضل؟ 25
س71/ لماذا تتطلب خوارزميات التعلم العميق كميات كبيرة من البيانات؟ 25
س72/ ما هو الفرق بين التكرار iteration والفترة epoch؟ 26
س73/ كيف تختار عدد الفترات epochs؟ 26
س74/ هل ستزيد الفترات epochs من الدقة accuracy؟ 26
س75/ ما حجم الدفعة batch size التي يجب أستخدامها في التعلم الآلي والعميق؟ 26
س76/ ما الفرق بين الحقبة epoch والدفعة batch والتكرار iteration؟ 26
س77/ هل حجم الدفعة الأكبر batch size أو الاصغر أفضل؟ 27
س78/ ما هو معدل التعلم learning rate الجيد؟ 27
س79/ هل معدل التعلم العالي جيد؟ 27
س80/ كيف تجد معدل التعلم المثالي؟ 27
س81/ هل زيادة معدل التعلم يؤثر على الدقة؟ 27
س82/ ما هو معدل التعلم المضمحل learning rate decay؟ 28
س83/ ما هو أفضل تقسيم لبيانات التدريب training والتحقق من الصحة validation والاختبار test؟ 28
س84/ ما هي النسبة المثلى لبيانات التدريب والاختبار؟ 28

س85/ ما هو استخدام ARIMA؟ 28
س86/ هل ARIMA تعلماً عميقاً؟ 28
س87/ ما هي مزايا نموذج ARIMA؟ 28
س88/ هل تستخدم Netflix التعلم العميق؟ 29
س89/ هل يستخدم Siri التعلم الآلي؟ 29
س90/ ما هو استخدام VGG16؟ 29
س91/ ما هو الفرق بين CNN وVGG16؟ 29
س92/ ما هي النماذج المدربة مسبقاً Pretrained model؟ 29
س93/ هل VGG16 نموذج مدرب مسبقًا Pretrained model؟ 29
س94/ كم عدد المعلمات parameters هل VGG16؟ 29
س95/ كيف نستخدم VGG16 في التعرف على الوجوه face recognition؟ 30
س96/ كيف يمكنني تحسين نموذج VGG16 الخاص بي؟ 30
س97/ كم عدد الخلايا العصبية في VGG؟ 30
س98/ كيف يمكنني تحسين دقة VGG16 الخاصة بي؟ 30
س99/ هل ResNet هي CNN؟ 31
س100/ كيف تختلف ResNet عن CNN؟ 31
س101/ ما هي معمارية ResNet50؟ 31
س102/ ما هو حجم الإدخال لـ ResNet50؟ 31
س103/ ما هو استخدام ResNet-50؟ 31
س104/ هل ResNet50 هو نقل التعلم transfer learning؟ 31
~~س105/ أيهما أفضل VGG16 أم ResNet؟~~ 32
س106/ لماذا تعتبر ResNet جيدة لتصنيف الصور؟ 32
~~س107/ لماذا تعتبر ResNet جيدة لتصنيف الصور؟~~ 32
س108/ ما هو نموذج Inception v3؟ 32
س109/ هل Inception v3 هو نموذج CNN؟ 32
س110/ كم عدد الطبقات الموجودة في Inception-v3؟ 32
س111/ ما هو الفرق بين inception v2 و v3؟ 33
س112/ هل Inception-v3 أفضل من VGG16؟ 33
س113/ ما هو حجم الإدخال input size في inception-v3؟ 33

س114/ هل RNN خاضعة للأشراف supervised ام غير خاضعة للأشراف unsupervised؟ 33
س115/ ما هي أنواع الـ RNN؟ 33
س116/ ما هي عيوب الـ RNN؟ 33
س117/ هل يمكننا استخدام RNN لتصنيف الصور؟ 34
س118/ ما نوع البيانات المناسبة لـ RNN؟ 34
س119/ كيف تختلف RNN عن الشبكات العصبية الأخرى؟ 34
س120/ هل يمكننا الجمع بين RNN وCNN؟ 34
س121/ هل CNN أسرع من RNN؟ 34
~~س122/ لماذا يتم استخدام الـ RNN للبيانات النصية؟~~ 34
س123/ هل يمكننا استخدام الشبكة العصبية المتكررة (RNN) للتصنيف الثنائي؟ 35
س124/ لماذا RNN أفضل من CNN لـ NLP؟ 35
س125/ ما هي المشاكل التي يمكن لـ RNN أفضل من الـ CNN؟ 35
س126/ ما هو الـ RNN في معالجة الصور؟ 35
س127/ كيف يتم تدريب RNN؟ 35
س128/ أيهما أفضل لتصنيف الصور CNN أم RNN؟ 35
س129/ ما هي الشبكة العصبية العميقة الأفضل لـ chatbot؟ 36
س130/ كيف يتم استخدام RNN في chatbot؟ 36
س131/ ما هو LSTM ولماذا يتم استخدامه؟ 36
س132/ لماذا LSTM أفضل من RNN؟ 36
س133/ كيف تعمل LSTM؟ 36
س134/ لماذا LSTM هو الأفضل للسلسلة الزمنية time series؟ 37
س135/ كم عدد الطبقات التي تمتلكها LSTM؟ 37
س136/ ما المشكلة التي تحلها LSTM؟ 37
س137/ لماذا يتم استخدام LSTM في RNN؟ 37
س138/ ما هي الاكثر العيوب التي تعاني منها الـ LSTM؟ 37
س139/ ما هي عيوب LSTM؟ 37
س140/ ما هي المعلمات parameters في LSTM؟ 38
س141/ ما هو الفرق بين CNN وLSTM؟ 38
س142/ هل يمكننا تطبيق نموذج LSTM لتصنيف الصور؟ 38

س143/هل LSTM جيدة للتصنيف؟ 38
س144/كيف يتجنب LSTM مشكلة تلاشي الانحدار؟ 38
س145/ما هي مدخلات ومخرجات LSTM؟ 38
س146/كيف تزيد الدقة accurecy في LSTM؟ 39
س147/لماذا يتم استخدام LSTM في البرمجة اللغوية العصبية NLP؟ 39
س148/كيف يعمل LSTM خطوة بخطوة؟ 39
س149/كيف يتم تدريب LSTM؟ 39
س150/ما هو استخدام GRU؟ 39
س151/كيف تختلف GRU عن RNN القياسية؟ 39
س152/أيهما أفضل LSTM أو GRU؟ 40
س153/كيف يمكنني إيقاف الضبط الزائد overfitting في LSTM؟ 40
س154/ما هو نموذج LSTM ثنائي الاتجاه Bi-LSTM؟ 40
س155/ما هو الفرق بين LSTM و Bi-LSTM؟ 40
س156/أيهما أفضل LSTM أم BiLSTM؟ 40
س157/ما هو استخدام LSTM ثنائي الاتجاه؟ 40
س158/لماذا LSTM بطيئة؟ 41
س159/هل تحتاج إلى GPU لـ LSTM؟ 41
س160/كم عدد طبقات LSTM التي أحتاجها؟ 41
س161/ماهي vanilla LSTM؟ 41
س162/لماذا LSTM أفضل من Arima؟ 41
س163/هل تدريب GRU أسهل من تدريب LSTM التقليدي؟ 41
س164/هل تمتلك GRU وحدة ذاكرة؟ 42
س165/كم عدد المعلمات parameters في GRU؟ 42
س166/كم عدد بوابات GRU؟ 42
س167/ما هو المحول transformer في الشبكة العصبية؟ 42
س168/ماهي استخدامات المحول transformer؟ 42
س169/ما هي المحولات في البرمجة اللغوية العصبية NLP؟ 42
س170/هل يمكن استخدام المحولات لتصنيف النص text classification؟ 42
س171/ما هي خوارزمية BERT؟ 43
س172/ما هي تطبيقات BERT؟ 43

س173/هل يمكن استخدام BERT للترجمة؟ 43
س174/هل يمكن أن تحل المحولات محل LSTM؟ 43
س175/لماذا المحولات transformers أفضل من LSTM؟ 43
س176/هل يعتمد BERT على المحولات؟ 43
س177/ما هو استخدام GAN؟ 44
س178/هل GAN لعبة مجموع صفري zero sum game؟ 44
س179/ما هي الأنواع المختلفة لشبكات GAN؟ 44
س180/ما هي التحديات التي تواجهها شبكات الخصومة التوليدية (GANs)؟ 44
س181/ما هي خطوات تنفيذ شبكة GAN؟ 45
س182/هل يمكن استخدام GAN للتصنيف؟ 45
س183/أي GAN هو الأفضل لتوليد الصور؟ 45
س184/هل يمكن استخدام GAN للانحدار regression؟ 45
س185/ما هو الفرق بين GAN وGAN الشرطي (C-GAN)؟ 46
س186/ما هو GAN ثنائي الاتجاه bidirectional GAN؟ 46
س187/لماذا يصعب تدريب GAN؟ 46
س188/هل GAN هي التزييف العميق Deepfake؟ 46
س189/ما هي عيوب GAN؟ 46
س190/كيف تزيد دقة GAN؟ 46
س191/لماذا يستخدم Tanh في المولدات في GAN؟ 47
س192/ما هو المحسن Optimizer الأفضل لـ GAN؟ 47
س193/ماذا يعني أن خسارة أداة التمييز بقيت عند قيمة ثابتة بينما تتناقص خسارة المولد؟ 47
س194/ما هو استخدام المشفر التلقائي autoencoder؟ 47
س195/هل يقلل المشفر التلقائي الأبعاد؟ 47
س196/هل المشفر التلقائي غير خطي nonlinear؟ 47
س197/هل المشفر التلقائي خاضع للإشراف أم غير خاضع للإشراف؟ 47
س198/ما هي الأنواع المختلفة للمشفر التلقائي؟ 48
س199/ما هو المشفر encoder في المشفر التلقائي Autoencoder؟ 48
س200/كيف تقلل المشفرات التلقائية الخطأ؟ 48
س201/ما هو عنق الزجاجة bottleneck في التشفير التلقائي؟ 48

س202/ ما هو الفضاء الكامنة latent space في المشفر التلقائي؟ 48
س203/ ما هي المشفرات التلقائية المتغيرة variational Autoencoder (VAE)؟ 48
س204/ ما هو الفرق بين المشفر التلقائي Autoencoder (AE) والمشفر التلقائي المتغير variational Autoencoder (VAE)؟ 49
س205/ لماذا نحتاج VAE؟ 49
س206/ هل المشفر التلقائي المتغير VAE أفضل من المشفر التلقائي AE؟ 49
س207/ ما هو العيب الأكثر شيوعاً في VAEs؟ 49
س208/ ما هو المشفر التلقائي التلافيفي convolutional autoencoder؟ 49
س209/ هل شبكات GAN أفضل من VAE؟ 49
س210/ هل يمكن استخدام المشفرات التلقائية لتقليل الضوضاء؟ 50
س211/ ما هو المشفر التلقائي لازالة الضوضاء Denoising Autoencoder؟ 50
س212/ ما هو الفرق الرئيسي بين المشفر التلقائي Autoencoder والمشفر التلقائي لإزالة الضوضاء Denoising Autoencoder؟ 50
س213/ ما هو المشفر التلقائي المتناثر sparse autoencoder ؟ 50
س214/ هل يمكن استخدام المشفرات التلقائي للتجميع clustering؟ 50
س215/ هل يمكن استخدام المشفرات التلقائي للتصنيف classification؟ 51
س216/ هل RTX 3080 جيد للتعلم العميق؟ 51
س217/ أي GPU هو الأفضل للتعلم العميق؟ 51
س218/ هل RTX أو GTX أفضل للتعلم العميق؟ 51
س219/ ما هي بطاقة الرسوميات CUDA؟ 51
س220/ ما هي نوى CUDA للتعلم العميق؟ 51
س221/ ما هو البيتافلوب petaflops؟ 52
س222/ ما هو CUDA في CNN؟ 52
س223/ ما هو التعلم المعزز reinforcement learning؟ 52
س224/ ما هي الحالة state والاجراء action في التعلم المعزز؟ 52
س225/ ما هي المكافأة reward والعقوبة penalty في التعلم المعزز؟ 52
س226/ ما هو المقصود بـ Q-Learning؟ 52
س227/ ما هي قيمة Q في التعلم المعزز؟ 53
س228/ ما الفرق بين Q-Learning و Sarsa؟ 53
س229/ ما هي أمثلة التعلم المعزز؟ reinforcement learning؟ 53

س230/هل التعلم المعزز جزء من الذكاء الاصطناعي AI؟ 53
س231/هل يحتاج التعلم المعزز إلى بيانات تدريبية trained data؟ 53
س232/ما هو الاختلاف الرئيسي بين التعلم العميق DL والتعلم المعزز RL؟..... 53
س233/ما هي عيوب التعلم المعزز؟ 54
س234/هل التعلم المعزز مستخدم في الصناعة industry؟ 54
س235/ما هي التطبيقات واقعية للتعلم المعزز..... 54
س236/ما هو المقصود بالتعلم المعزز العميق deep reinforcement learning؟ 54
س237/ما هو التعلم المعزز العميق وما هي تطبيقاته المختلفة؟..... 55
س238/ما هو الفرق بين التعلم المعزز والتعلم المعزز العميق؟ 55
س239/ما هو التعلم العميق الهندسي Geometric Deep Learning؟ 55
س240/ما هو التعلم العميقQ (Deep Q Learning)؟..... 55
س241/ما هو التعلم الانتقالي العميق Deep transfer learning (DTL)؟ 55
س242/ما هو التعلم الشامل end to end learning في التعلم العميق؟ 56
س243/ماذا تعني قاعدة دلتا Delta Rule في التعلم الآلي والعميق؟ 56
س244/ما الخطوة التالية للتعلم الآلي؟ 56
س245/ما الحوسبة الكمومية Quantum computing؟ 56
س246/هل يمكن استخدام أجهزة الكمبيوتر الكمومية للتعلم العميق؟..... 56
س247/هل تعلم الآلة الكمي quantum machine learning حقيقي؟..... 56
س248/ما المقصود بالتعلم الآلي الكمي quantum machine learning ؟ 56
س249/ما هو الجيل القادم من الذكاء الاصطناعي؟..... 57
س250/ما هي الأنواع الثلاثة للذكاء الاصطناعي؟..... 57
س251/ما الذي يرمز إليه الذكاء الاصطناعي العام GAI في الذكاء الاصطناعي؟..... 57
س252/ ما الفرق بين الذكاء الاصطناعي الضيق Narrow AI والذكاء الاصطناعي العام General AI؟..... 57
س253/ما هو الديب مايند DeepMind؟..... 57
س254/ما هي البرامج التي يستخدمها DeepMind؟ 58
س255/لماذا يحظى التعلم العميق بشعبية في السنوات الأخيرة؟..... 58
س256/ما هو مستقبل التعلم العميق؟..... 58

سؤال

هل يمكن استخدام أجهزة الكمبيوتر الكمومية للتعلم العميق؟

س1/ ما هو نقل التعلم Transfer Learning في التعلم العميق؟

ج/ تُعرف إعادة استخدام نموذج تم تعلمه مسبقًا في مشكلة جديدة باسم نقل التعلم **Transfer Learning**. إنه شائع بشكل خاص في التعلم العميق في الوقت الحالي لأنه يمكنه تدريب الشبكات العصبية العميقة بكمية صغيرة من البيانات.

س2/ ما هو نقل التعلم لـ CNN؟

ج/ الفرضية الأساسية لتعلم النقل بسيطة: خذ نموذجًا تم تدريبه على مجموعة بيانات كبيرة وانقل معرفته إلى مجموعة بيانات أصغر. للتعرف على الكائنات باستخدام **CNN**، نقوم بتجميد الطبقات التلافيفية **Convolution Layer** المبكرة للشبكة وندرب فقط الطبقات القليلة الأخيرة التي تقوم بالتنبؤ.

س3/ ما هو نقل التعلم مع المثال المناسب؟

ج/ في نقل التعلم، تستغل الآلة المعرفة **knowledge** المكتسبة من مهمة سابقة لتحسين التعميم حول مهمة أخرى. على سبيل المثال، في تدريب المصنف على التنبؤ بما إذا كانت الصورة تحتوي على طعام، يمكنك استخدام المعرفة التي اكتسبتها أثناء التدريب للتعرف على المشروبات.

س4/ لماذا نحتاج نقل التعلم؟

ج/ يساعد التعلم الانتقالي المطورين على اتباع نهج مدمج من نماذج مختلفة لضبط حل لمشكلة معينة. يمكن أن يؤدي تبادل المعرفة بين نموذجين مختلفين إلى نموذج أكثر دقة وقوة. النهج يسمح لنماذج البناء بطريقة تكرارية.

س5/ ما هو الفرق بين نقل التعلم transfer learning والضبط الدقيق fine-tuning؟

ج/ نقل التعلم هو عندما يتم إعادة استخدام نموذج تم تطويره لمهمة واحدة للعمل في مهمة ثانية. يعد الضبط الدقيق أحد الأساليب لنقل التعلم حيث تقوم بتغيير إخراج النموذج ليلائم المهمة الجديدة وتدريب نموذج الإخراج فقط.

س6/ ما هو الفرق بين الشبكة العصبية التلافيفية CNN و نقل التعلم Transfer Learning؟

ج/ بالنسبة إلى **CNN**، تحتاج إلى مزيد من المعالجة المسبقة لمجموعة البيانات ولكن مع نقل التعلم، لا تحتاج إلا إلى القليل من معالجة مجموعة البيانات مثل تغيير الحجم إلى **227 × 227** أو **224 × 224** وفقًا للنماذج المحددة مسبقًا (**AlexNet** و

GoogLeNet و **ResNet** و **VGG Networks** وما إلى ذلك. المزيد) هذا يوفر الكثير من وقت المعالجة المسبقة للبيانات.

س7/هل نقل التعلم دائمًا أفضل؟

ج/بشكل عام، عند استخدامه بشكل مناسب، سيمنحك تعليم النقل ثلاث فوائد: دقة بداية أعلى وتقارب أسرع ودقة تقارب أعلى (مستوى الدقة الذي يتقارب معه التدريب).

س8/ماذا نعني بتجميد الطبقات freeze layers؟

ج/تجميد طبقة يمنع تعديل أوزانها. غالبًا ما تُستخدم هذه التقنية في نقل التعلم، حيث يتم تجميد النموذج الأساسي (المدرب على مجموعة بيانات أخرى).

س9/في أي الحالات يمكننا استخدام نقل التعلم؟

ج/ يُعد نقل التعلم مفيدًا جدًا بشكل خاص عندما يكون لديك مجموعة بيانات تدريب صغيرة. في هذه الحالة، يمكنك، على سبيل المثال، استخدام الأوزان من النماذج المدربة مسبقًا لتهيئة أوزان النموذج الجديد. كما سترى لاحقًا، يمكن أيضًا تطبيق نقل التعلم على مشاكل معالجة اللغة الطبيعية **NLP**.

س10/ما هي عيوب نقل التعلم؟

ج/ حاليًا، إحدى أكبر القيود على نقل التعلم هي مشكلة النقل السلبي. لا يعمل التعلم الانتقالي إلا إذا كانت المشكلات الأولية والهدف متشابهة بدرجة كافية حتى تكون الجولة الأولى من التدريب ذات صلة.

س11/ما فائدة نقل التعلم؟

ج/تشمل الفوائد الرئيسية لتعلم النقل توفير الموارد وتحسين الكفاءة عند تدريب نماذج جديدة. يمكن أن يساعد أيضًا في نماذج التدريب عند توفر مجموعات البيانات غير الموسومة فقط، حيث سيتم تدريب الجزء الأكبر من النموذج مسبقًا.

س12/ما هي مبادئ نقل التعلم؟

ج/في نقل التعلم، نقوم أولاً بتدريب شبكة أساسية على مجموعة بيانات ومهمة أساسية، ثم نقوم بإعادة توظيف الميزات المكتسبة، أو نقلها، إلى شبكة هدف ثانية ليتم تدريبها على مجموعة بيانات ومهمة مستهدفة.

س13/ ما هي الشبكة العصبية التلافيفية CNN؟

ج/ ضمن التعلم العميق، تعد الشبكة العصبية التلافيفية أو **CNN** نوعًا من الشبكات العصبية الاصطناعية **ANN**، والتي تُستخدم على نطاق واسع للتعرف على الصور / الكائنات وتصنيفها. وبالتالي يتعرف التعلم العميق على الكائنات في الصورة باستخدام **CNN**.

س14/ هل CNN خوارزمية أم نموذج؟

ج/ الشبكة العصبية التلافيفية (**ConvNet / CNN**) هي خوارزمية التعلم العميق التي يمكن أن تأخذ صورة إدخال ، وتعيين الأهمية (الأوزان القابلة للتعلم والتحيزات) للجوانب/الكائنات المختلفة في الصورة وتكون قادرة على تمييز أحدهما عن الآخر.

س15/ هل CNN خاضعة للإشراف supervised أم غير خاضعة للإشراف unsupervised؟

ج/ **CNN** هو نوع خاضع للإشراف من التعلم العميق، ويفضل استخدامه في التعرف على الصور **image recognition** ورؤية الكمبيوتر **computer vision**.

س16/ هل تستخدم CNN للصور فقط؟

ج/ كلا، لكن يفضل استخدامه في التعرف على الصور ورؤية الكمبيوتر.

س17/ هل يمكن استخدام CNN للبيانات الرقمية numerical data؟

ج/ نعم، يمكنك استخدام شبكة **CNN**. لا تقتصر قنوات **CNN** على الصور فقط. استخدم التفافًا أحادي الأبعاد **1D**، وليس التفافًا ثنائي الأبعاد؛ لديك بيانات **1D**، لذا فإن الالتفاف **1D** هو الأنسب.

س18/ هل يمكن استخدام CNN لغير الصور non-image ؟

ج/ كانت المحاولات السابقة لتطبيق **CNN** على البيانات غير المصورة مقتصرة على معماريات **1D CNN**. تكون المدخلات إلى **1D CNN** في شكل متجهات ميزة **feature vectors** وبالتالي لا يمكنها التعامل مع الصور.

س19/ لماذا تعد CNN أفضل من KNN؟

ج/ **KNN** و**CNN** يعملان بشكل تنافسي مع الخوارزمية الخاصة بهما في مجموعة البيانات **dataset** ، لكن **CNN** تنتج دقة عالية من **KNN** وبالتالي تم اختيارها كطريقة أفضل.

س20/ كيف يمكنني برمجة CNN من البداية؟

ج/ سنبني شبكة عصبية تلافيفية **CNN**، ثم ندربها ونختبرها. لذا، للبدء، نحتاج إلى المضي قدمًا خطوة بخطوة بطريقة هرمية.

1. إعداد بيانات التدريب والاختبار.
2. قم ببناء طبقات **CNN** باستخدام مكتبة **Tensorflow**.
3. حدد المُحسِّن **Optimizer**.
4. تدريب الشبكة وحفظ **checkpoints**.
5. أخيرًا ، نقوم باختبار النموذج.

س21/ هل يمكننا استخدام CNN لتصنيف النص text classification؟

ج/ نعم يمكننا استخدام **CNN** لتنصنيف النص. هناك العديد من الطرق لأداء تصنيف النص. **TextCNN** هي أيضًا طريقة تشير إلى الشبكات العصبية التلافيفية لإجراء تصنيف النص.

س22/ هل يمكننا استخدام CNN لتحليل المشاعر sentiment analysis؟

ج/ نعم يمكننا استخدام **CNN** لتحليل المشاعر. على سبيل المثال استخدام الشبكات العصبية التلافيفية **CNN** لتحليل المشاعر في مجموعة بيانات **IMDb**.

س23/ لماذا تعتبر CNN أفضل لتصنيف الصور؟

ج/ الميزة الرئيسية لـ **CNN** هي أنها تكتشف تلقائيًا الميزات المهمة دون أي إشراف بشري. هذا هو السبب في أن **CNN** ستكون حلاً مثاليًا لمشكلات رؤية الكمبيوتر وتصنيف الصور.

س24/ هل CNN أفضل من SVM لتصنيف الصور؟

ج/ من الواضح أن **CNN** تتفوق على مصنف **SVM** من حيث دقة الاختبار **testing accuracy**. عند مقارنة التصحيحات الإجمالية لمصنف **CNN** و **SVM**، تم تحديد **CNN** تتمتع بميزة ثابتة على **SVM** عند استخدام عينات الانعكاس المستندة إلى البكسل، دون حجم التجزئة **segmentation size**.

س25/ ما هي تطبيقات CNN؟

ج/ لدى **CNN** تطبيقات في التعرف على الصور والفيديو، وأنظمة التوصية، وتصنيف الصور، وتجزئة الصور، وتحليل الصور الطبية، ومعالجة اللغة الطبيعية، والرؤية الحاسوبية، والسلاسل الزمنية المالية.

س26/ هل يمكنني استخدام CNN في السلاسل الزمنية time series؟

ج/ على الرغم من شيوع شبكة **CNN** في مجموعات بيانات الصور، إلا أنه يمكن استخدامها (وقد تكون أكثر عملية من **RNNs**) في بيانات السلاسل الزمنية. قدم بنية شائعة لتصنيف السلاسل الزمنية (متغير ومتعدد المتغيرات) تسمى الشبكة العصبية التلافيفية الكاملة (**FCN**).

س27/ هل رؤية الكمبيوتر هي نفسها CNN؟

ج/ كلا، لكن تستخدم معظم خوارزميات الرؤية الحاسوبية شيئًا يسمى الشبكة العصبية التلافيفية أو **CNN. CNN** هو نموذج يستخدم في التعلم الآلي لاستخراج الميزات ، مثل النسيج والحواف ، من البيانات المكانية **spatial data**.

س28/ هل Yolo مشتق من CNN؟

ج/ تستخدم خوارزمية **YOLO** الشبكات العصبية التلافيفية (**CNN**) لاكتشاف الكائنات في الوقت الفعلي. كما يوحي الاسم ، لا تتطلب الخوارزمية سوى انتشار أمامي واحد **single forward propagation** عبر شبكة عصبية لاكتشاف الكائنات.

س29/ هل CNN مناسبة لـ NLP؟

ج/ يمكن استخدام شبكات **CNN** في مهام التصنيف المختلفة في البرمجة اللغوية العصبية **NLP**. الالتفاف **convolution** هو نافذة تنزلق فوق بيانات إدخال أكبر مع التركيز على مجموعة فرعية من مصفوفة الإدخال **input matrix**. يعد الحصول على بياناتك بالأبعاد الصحيحة أمرًا بالغ الأهمية لأي خوارزمية تعلم.

س30/ هل يمكن استخدام CNN لاختيار الميزة feature selection؟

ج/ يعد اختيار الميزة أسلوبًا مهمًا لتحسين أداء الشبكة العصبية نظرًا للسمات الزائدة عن الحاجة والكمية الهائلة في مجموعات البيانات الأصلية.

س31/ هل يمكننا استخدام CNN لاستخراج الميزة؟

ج/ تستخدم **CNN** كمستخرج الميزة **feature extractor** في عملية التدريب بدلاً من تنفيذها يدويًا. يتكون مستخرج ميزة **CNN** من أنواع خاصة من الشبكات العصبية التي تحدد الأوزان من خلال عملية التدريب.

س32/ ما هي فائدة CNN؟

ج/ تُستخدم شبكات **CNN** بشكل أساسي لتصنيف الصور والتعرف عليها. إن تخصص **CNN** هو قدرتها التلافيفية **convolutional ability**. إن إمكانية استخدامات أخرى

لشبكات **CNN** لا حدود لها وتحتاج إلى استكشافها ودفعها إلى حدود أخرى لاكتشاف كل ما يمكن تحقيقه بواسطة هذه الآلية المعقدة.

س33/ لماذا يستخدم Max pooling في CNN؟

ج/ يساعد التجميع **pooling** بشكل أساسي في استخراج ميزات حادة **sharp** وسلسة **smooth**. يتم إجراء ذلك أيضًا لتقليل التباين والحسابات. يساعد **Max-pooling** في استخراج ميزات منخفضة المستوى مثل الحواف **edges** والنقاط **points** وما إلى ذلك. بينما يذهب **Avg-pooling** إلى ميزات سلسة **smooth features**.

س34/ لماذا نفضل CNN على ANN بالنسبة لبيانات الصورة كمدخلات؟

ج/ نظرًا لأن الصور الرقمية عبارة عن مجموعة من وحدات البكسل ذات القيم العالية، فمن المنطقي استخدام **CNN** لتحليلها. تقلل **CNN** من قيمها، وهو أفضل لمرحلة التدريب مع قوة حسابية أقل وفقدان أقل للمعلومات.

س35/ كيف يمكنني إيقاف الضبط الزائد overfitting في CNN؟

ج/ خطوات لتقليل فرط التعلم (الضبط الزائد) في **CNN**:

1. أضف المزيد من البيانات.
2. استخدام زيادة البيانات **data augmentation**.
3. استخدم البنى التي تُعمم **generalize** جيدًا.
4. أضف التنظيم **regularization** (في الغالب الحذف العشوائي **dropout** ، من الممكن أيضًا تسوية **L1 / L2**)
5. تقليل تعقيد الهيكلية.

س36/ هل يقلل التجميع pooling من فرط التعلم overfitting في CNN؟

ج/ يوفر التجميع القدرة على تعلم الميزات الثابتة ويعمل أيضًا كمنظم **regularizer** لتقليل مشكلة **overfitting**. بالإضافة إلى ذلك، فإن تقنيات التجميع تقلل بشكل كبير من التكلفة الحسابية ووقت التدريب للشبكات التي لا تقل أهمية عن وضعها في الاعتبار.

س37/ ماذا تفعل R-CNN؟

ج/ **R-CNN** هي شبكة تلافيفية عميقة تستخدم لاكتشاف الأشياء **object detection**، والتي تظهر للمستخدم كشبكة واحدة، من طرف إلى طرف، وموحدة. يمكن للشبكة أن تتنبأ بدقة وبسرعة بمواقع كائنات مختلفة.

س38/ هل R-CNN تعلماً عميقاً؟

ج/ نعم، المناطق ذات الشبكات العصبية التلافيفية **(R-CNN)**، يجمع بين مقترحات المنطقة المستطيلة **rectangular region proposals** مع ميزات الشبكة العصبية التلافيفية.

س39/ ما هو المحسن Optimizer الأفضل لشبكة CNN؟

~~ج/آدم **Adam** هو أفضل المحسنين.~~

س40/ ما هو معدل التعلم learning rate في CNN؟

ج/ يشار إلى المقدار الذي يتم تحديث الأوزان أثناء التدريب على أنه حجم الخطوة أو "معدل التعلم". على وجه التحديد، معدل التعلم عبارة عن معلمة فائقة **hyperparameter** قابلة للتكوين تستخدم في تدريب الشبكات العصبية التي لها قيمة إيجابية صغيرة، غالبًا في النطاق بين **0.0** و **1.0**.

س41/ ما هي أفضل طريقة لاختيار معمارية CNN؟

ج/

- التجريب **Experimentation**: جرب عددًا مختلفًا من الطبقات والعقد.
- الحدس **Intuition**: استخدم الخبرة السابقة لاختيار عدد الطبقات والعقد.
- ابحث عن العمق **Go for depth**: غالبًا ما تؤدي الشبكات العصبية العميقة أداءً أفضل من الشبكات الضحلة.
- استعارة الأفكار **Borrow ideas**: استعارة الأفكار من المقالات التي تصف مشاريع مماثلة.

س42/ ما هو نموذج CNN الأفضل لتصنيف الصور image classification؟

ج/ **VGG16** هو نموذج **CNN** مدرب مسبقًا ويستخدم لتصنيف الصور. يتم تدريبه على مجموعة بيانات كبيرة ومتنوعة ويتم ضبطه بدقة لملاءمة مجموعات بيانات تصنيف الصور بسهولة.

س43/ ما هو استخدام U-Net؟

ج/ تم اختراع **U-net** في الأصل واستخدم لأول مرة لتجزئة الصور الطبية الحيوية **biomedical image segmentation**. يمكن اعتبار بنيتها على نطاق واسع على أنها شبكة تشفير **encoder network** متبوعة بشبكة مفكك تشفير **decoder network**.

س44/ هل U-Net هو نوع من CNN؟

ج/ **UNet** هي بنية شبكة عصبية تلافيفية **CNN** توسعت مع تغييرات قليلة في بنية **CNN**. تم اختراعه للتعامل مع الصور الطبية الحيوية **biomedical images** حيث يكون الهدف ليس فقط تصنيف ما إذا كانت هناك عدوى أم لا ولكن أيضًا لتحديد منطقة الإصابة.

س45/ هل U-Net تلافيفية بالكامل fully convolutional؟

ج/ إنها شبكة عصبية تلافيفية بالكامل مصممة للتعلم من عدد أقل من عينات التدريب.

س46/ ما هو اكثر نماذج التعلم العميق للتعرف على الوجوه face recognition شيوعاً؟

ج/ أكثر نماذج التعرف على الوجوه شيوعًا:

- VGG-Face
- Google FaceNet
- OpenFace
- Facebook DeepFace
- DeepID
- Dlib
- ArcFace

س47/ ما هو Deepface؟

ج/ **DeepFace** هي مكتبة التعرف على الوجوه وتحليل سمات الوجه الأكثر وزنًا في **Python**. تتضمن مكتبة **DeepFace** مفتوحة المصدر جميع نماذج الذكاء الاصطناعي الرائدة للتعرف على الوجوه وتعالج تلقائيًا جميع إجراءات التعرف على الوجه في الخلفية.

س48/ ما هو DeepDream؟

ج/ **DeepDream** هو برنامج رؤية كمبيوتر تم إنشاؤه بواسطة مهندس **Google Alexander Mordvintsev** والذي يستخدم شبكة عصبية تلافيفية للعثور على أنماط في الصور وتحسينها عبر خوارزمية **pareidolia** ، وبالتالي خلق مظهر يشبه الحلم يذكرنا بتجربة مخدرة في الصور المعالجة بشكل مفرط عن عمد.

س49/ كيف يخلق الذكاء الاصطناعي الفن؟

ج/ لإنشاء فن الذكاء الاصطناعي، يكتب الفنانون خوارزميات ليس لاتباع مجموعة من القواعد، ولكن "لتعلم" جمالية معينة **specific aesthetic** من خلال تحليل آلاف الصور **textual descriptions**. ثم تحاول الخوارزمية بعد ذلك إنشاء صور جديدة بالالتزام بالجماليات **aesthetics** التي تعلمتها.

س50/ ما هو Midjourney؟

ج/ **Midjourney** هو مختبر أبحاث واسم برنامج الذكاء الاصطناعي للمختبر الذي ينشئ صورًا من الأوصاف النصية، على غرار برنامج **DALL-E** الخاص بـ **OpenAI**. الأداة حاليًا في إصدار تجريبي مفتوح..

س51/ لماذا تعتبر CNN الأفضل للتعرف على الوجوه؟

ج/ تمت دراسة بناء وتدريب نموذج **CNN** على أساس التعرف على الوجوه. لتبسيط نموذج **CNN**، يتم دمج طبقات الالتفاف **convolution** وأخذ العينات **sampling** في طبقة واحدة. استنادًا إلى الشبكة المدربة بالفعل، قم بتحسين معدل التعرف على الصور بشكل كبير.

~~س52/ لماذا تعتبر CNN الأفضل للتعرف على الوجوه؟~~

~~ج/ تمت دراسة بناء وتدريب نموذج **CNN** على أساس التعرف على الوجوه. لتبسيط نموذج **CNN**، يتم دمج طبقات الالتفاف **convolution** وأخذ العينات **sampling** في طبقة واحدة. استنادًا إلى الشبكة المدربة بالفعل، قم بتحسين معدل التعرف على الصور بشكل كبير.~~

س53/ ما الخوارزمية المستخدمة في التعرف على الوجوه في التعلم الآلي ML؟

ج/ هناك خوارزمية مستخدمة على نطاق واسع تسمى خوارزمية **Viola-Jones**. تعد خوارزمية **Viola-Jones** مرغوبة للغاية نظرًا لمعدل الكشف العالي ووقت المعالجة السريع.

س54/ ما هو نموذج CNN المدرب مسبقاً pre trained model الأفضل للتعرف على الوجوه؟

ج/ **VGGFace** هو نموذج مدرب مسبقًا يستخدم في أنظمة التعرف على الوجه للتعرف على بطاقات الهوية الإلكترونية التي يسببها التطور العمري **age development**. يستخدم **VGGFace2** بنية **ResNet-50**.

س55/ ما الخوارزمية المستخدمة للتعرف على الوجوه في بايثون؟

ج/ يوفر OpenCV بعض خوارزميات التعرف على الوجه التقليدية.

~~س56/ ما الخوارزمية المستخدمة للتعرف على الوجوه في بايثون؟~~

~~ج/ يوفر OpenCV بعض خوارزميات التعرف على الوجه التقليدية.~~

س57/ لماذا تعتبر ResNet جيدة لتصنيف الصور؟

ج/ يتكون **ResNet-50** من عمق **50** طبقة ويتم تدريبه على مليون صورة من **1000** فئة من قاعدة بيانات **ImageNet**. علاوة على ذلك، يحتوي النموذج على أكثر من **23** مليون معلمة قابلة للتدريب، مما يشير إلى بنية عميقة تجعله أفضل للتعرف على الصور.

س58/ ما هي لغة البرمجة الأفضل للتعرف على الوجه Face recognition؟

ج/ أفضل لغات البرمجة للتعرف على الوجوه:

- **OpenCV**
- ماتلاب
- **C / C ++ / C #**
- بايثون
- جافا

س59/ ما هي الطبقة الكثيفة Dense Layer في CNN؟

ج/ الطبقة الكثيفة **Dense Layer** هي طبقة بسيطة من الخلايا العصبية حيث يتلقى كل خلية عصبية مدخلات من جميع الخلايا العصبية في الطبقة السابقة، وبالتالي تسمى كثيفة **Dense Layer**. تُستخدم الطبقة الكثيفة لتصنيف الصورة بناءً على الإخراج من الطبقات التلافيفية **Convolution Layers**. تحتوي هذه الطبقة على عدد متعدد من هذه الخلايا العصبية.

س60/ ما هي طبقة Softmax في CNN؟

ج/ تُستخدم دالة **softmax** كدالة تنشيط في طبقة الإخراج لنماذج الشبكة العصبية التي تتنبأ بتوزيع احتمالي متعدد الحدود. أي، يتم استخدام **softmax** كدالة تنشيط لمشاكل التصنيف متعدد الفئات حيث تكون عضوية الفئة مطلوبة على أكثر من تسميتين (علامتين) للفئة.

~~س61/ لماذا نقوم بتسوية البيانات data normalization في CNN؟~~

~~ج/ بشكل عام، السبب وراء تسوية البيانات هو جعل النموذج يتقارب بشكل أسرع. عندما لا يتم تسوية البيانات، فإن الأوزان المشتركة للشبكة لها معايرات مختلفة لميزات مختلفة، مما يجعل دالة الخطأ cost function تتقارب ببطء شديد وغير فعالة.~~

س62/ لماذا يتم استخدام ReLU في الطبقات المخفية؟

ج/ ربما تكون دالة التنشيط الخطي المصححة **rectified linear activation**، أو دالة التنشيط **ReLU** ، هي الدالة الأكثر شيوعًا المستخدمة للطبقات المخفية. إنه شائع لأنه سهل التنفيذ وفعال في التغلب على قيود دوال التنشيط الأخرى الشائعة سابقًا، مثل **Sigmoid** و**Tanh**. يتقارب النموذج المدرب مع **ReLU** بسرعة وبالتالي يستغرق وقتًا أقل بكثير مقارنة بالنماذج المدربة على دالة **Sigmoid** و **Tanh**. يمكننا أن نرى بوضوح الضبط الزائد **overfitting** في النموذج المدرب مع **ReLU**. هذا بسبب التقارب السريع. يكون أداء النموذج أفضل بشكل ملحوظ عند التدريب مع **ReLU**

س63/ هل CNN هو الأفضل لاستخراج الميزات؟

ج/ نعم، شبكات **CNN** توفر طريقة استخراج ميزات تلقائية ، وهي ميزة أساسية ومهمة.

س64/ هل تحتاج CNN إلى اختيار الميزة feature selection؟

ج/ نعم، تحتاج ذلك. يعد اختيار الميزة **feature selection** أسلوبًا مهمًا لتحسين أداء الشبكة العصبية بسبب السمات الزائدة والكم الهائل في مجموعات البيانات الأصلية.

س65/ ما هو الفرق بين الطبقة التلافيفية convolution layer وطبقة كثيفة dense layer؟

ج/ الاختلاف الرئيسي بين الطبقة التلافيفية والطبقة الكثيفة هو أن الطبقة التلافيفية تستخدم عددًا أقل من المعلمات عن طريق إجبار قيم الإدخال على مشاركة المعلمات. تستخدم الطبقة الكثيفة عملية خطية بمعنى أن كل ناتج يتم تشكيله بواسطة الدالة بناءً على كل إدخال.

س66/ ما المقصود بالفترة epoch في CNN؟

ج/ الفترة **epoch** هي مرة واحدة تتم معالجة جميع الصور مرة واحدة بشكل فردي من الأمام والخلف إلى الشبكة، فهذه فترة واحدة **one epoch**. يتم احتساب الفترة

الواحدة عند (عدد التكرارات * حجم الدفعة) / العدد الإجمالي للصور في التدريب.
الفترة تعني تدريب الشبكة العصبية بجميع بيانات التدريب لدورة واحدة. في فترة ما، نستخدم جميع البيانات مرة واحدة بالضبط. يتم احتساب الممر الأمامي والممر الخلفي معًا كمرور واحد: تتكون الفترة من دفعة **batch** واحدة أو أكثر، حيث نستخدم جزءاً من مجموعة البيانات لتدريب الشبكة العصبية.

س67/ هل softmax دالة تنشيط activation function ام دالة خسارة (خطأ) loss function؟

ج/ تُستخدم دالة **softmax** كدالة تنشيط في طبقة الإخراج لنماذج الشبكة العصبية التي تتنبأ بتوزيع احتمالي متعدد الحدود. أي، يتم استخدام **softmax** كدالة تنشيط لمشاكل التصنيف متعدد الفئات حيث تكون عضوية الفئة مطلوبة على أكثر من علامتين (تسميتين) للفئة.

س68/ ما هو حجم الدفعة batch size في CNN؟

ج/ حجم الدُفعة هو معلمة فائقة **hyperparameter** للتدرج الاشتقاقي الذي يتحكم في عدد عينات التدريب للعمل من خلالها قبل تحديث المعلمات الداخلية للنموذج.

س69/ لماذا نحتاج الفترة epoch في التعلم الآلي والعميق؟

ج/ فترة واحدة تتكون من العديد من خطوات تحديث الوزن. فترة واحدة تعني أن المُحسِّن **optimizer** استخدم كل مثال تدريب مرة واحدة.لذلك تلعب الفترة **Epoch** دورًا مهمًا في نمذجة التعلم الآلي والعميق ، حيث أن هذه القيمة هي المفتاح للعثور على النموذج الذي يمثل العينة بأقل خطأ.

س70/ هل المزيد من الفترات more epochs أفضل؟

ج/ سيؤدي زيادة عدد الفترات إلى الضبط الزائد **overfitting** بينما سيؤدي تقليل عدد الفترات إلى الضبط الناقص **underfitting**. يمكنك اختيار التوقف المبكر **EarlyStopping** في **keras** الذي سيتوقف نموذج التدريب في فترة معينة بمجرد توقف أداء النموذج عن التحسن.

س71/ لماذا تتطلب خوارزميات التعلم العميق كميات كبيرة من البيانات؟

ج/ ستحتاج النماذج الأكثر قوة ودقة إلى مزيد من المعلمات، والتي بدورها تتطلب المزيد من البيانات.

س72/ ما هو الفرق بين التكرار iteration والفترة epoch؟

ج/ التكرار **iteration** هو معالجة لمرة واحدة للأمام وللخلف لمجموعة من الصور (لنفترض أن دفعة واحدة تم تعريفها على أنها **16**، ثم تتم معالجة **16** صورة في تكرار واحد). الفترة **epoch** هو مرة واحدة تتم معالجة جميع الصور مرة واحدة بشكل فردي من الأمام والخلف إلى الشبكة، فهذه فترة واحدة.

س73/ كيف تختار عدد الفترات epochs؟

ج/ يعتمد العدد الصحيح للفترات على الارتباك المتأصل (أو التعقيد) لمجموعة البيانات الخاصة بك. من القواعد الأساسية الجيدة أن تبدأ بقيمة تساوي **3** أضعاف عدد الأعمدة في بياناتك. إذا وجدت أن النموذج لا يزال يتحسن بعد اكتمال كل الفترات، فحاول مرة أخرى بقيمة أعلى.

س74/ هل ستزيد الفترات epochs من الدقة accuracy؟

ج/ لا تكون زيادة الفترات منطقية إلا إذا كان لديك الكثير من البيانات في مجموعة البيانات الخاصة بك. ومع ذلك، سيصل نموذجك في النهاية إلى نقطة لا تؤدي فيها الفترات المتزايدة إلى تحسين الدقة.

س75/ ما حجم الدفعة batch size التي يجب أستخدامها في التعلم الآلي والعميق؟

ج/ من الناحية العملية، لتحديد حجم الدُفعة الأمثل، نوصي بتجربة أحجام الدُفعات الأصغر أولاً (عادةً **32** أو **64**)، مع الأخذ في الاعتبار أيضًا أن أحجام الدُفعات الصغيرة تتطلب معدلات تعلم صغيرة. يجب أن يكون عدد أحجام الدُفعات بقوة **2** للاستفادة الكاملة من معالجة وحدات معالجة الرسوميات.

س76/ ما الفرق بين الحقبة epoch والدفعة batch والتكرار iteration؟

ج/

الفترة **epoch**: يمثل تكرارًا واحدًا على مجموعة البيانات بأكملها (كل شيء يتم وضعه في نموذج التدريب).

الدفعة **batch**: تشير إلى الوقت الذي لا يمكننا فيه تمرير مجموعة البيانات بأكملها إلى الشبكة العصبية في وقت واحد ، لذلك نقسم مجموعة البيانات إلى عدة دفعات.

التكرار **iteration**: إذا كان لدينا **10000** صورة كبيانات وحجم دفعة **200**. فيجب أن يتم تشغيل **50** مرة (**10000** مقسومة على **50**).

س77/ هل حجم الدفعة الأكبر batch size أو الاصغر أفضل؟

ج/ لا يؤدي حجم الدفعة الأصغر (ليس صغيرًا جدًا) عادةً إلى عدد أقل من التكرارات لخوارزمية التدريب، بدلاً من حجم الدفعة الكبير، ولكن أيضًا إلى دقة أعلى بشكل عام، أي، شبكة عصبية تعمل بشكل أفضل ، في نفس القدر من وقت التدريب ، أو أقل. تُستخدم أحجام الدُفعات الأصغر لسببين رئيسيين:

- أحجام الدُفعات الأصغر صاخبة **noisy**، وتوفر تأثيرًا منتظمًا وخطأ أقل في التعميم.
- تسهل أحجام الدُفعات الأصغر ملاءمة مجموعة واحدة من بيانات التدريب في الذاكرة (أي عند استخدام وحدة معالجة الرسومات **GPU**).

س78/ ما هو معدل التعلم learning rate الجيد؟

ج/ نطاق القيم التي يجب مراعاتها لمعدل التعلم أقل من **1.0** وأكبر من **10^-6**. القيمة الافتراضية التقليدية لمعدل التعلم هي **0.1** أو **0.01** ، وقد يمثل هذا نقطة بداية جيدة لمشكلتك.

س79/ هل معدل التعلم العالي جيد؟

ج/ يمكن أن يتسبب معدل التعلم الكبير جدًا في تقارب النموذج بسرعة كبيرة مع حل دون المستوى الأمثل **suboptimal solution** ، في حين أن معدل التعلم الصغير جدًا يمكن أن يتسبب في توقف العملية.

س80/ كيف تجد معدل التعلم المثالي؟

ج/ هناك عدة طرق لتحديد نقطة بداية جيدة لمعدل التعلم. تتمثل الطريقة الشائعة في تجربة بعض القيم المختلفة ومعرفة أي منها يمنحك اقل خطأ دون التضحية بسرعة التدريب. قد نبدأ بقيمة كبيرة مثل **0.1**، ثم نجرب قيمًا أقل أضعافًا مضاعفة: **0.01** ، **0.001** ، إلخ.

س81/ هل زيادة معدل التعلم يؤثر على الدقة؟

ج/ كلا. يؤثر معدل التعلم على السرعة التي يمكن أن يتقارب بها نموذجنا مع الحد الأدنى المحلي **local minima** (ويعرف أيضًا باسم الوصول إلى أفضل دقة). وبالتالي، فإن القيام بذلك بشكل صحيح من البداية يعني وقتًا أقل لتدريب النموذج.

س82/ ما هو معدل التعلم المضمحل learning rate decay؟

ج/ إنها تقنية لتدريب الشبكات العصبية الحديثة. يبدأ بمعدل تعلم كبير ثم يتلاشى عدة مرات. انه يساعد في كل من التحسين **optimization** والتعميم **generalization**.

س83/ ما هو أفضل تقسيم لبيانات التدريب training والتحقق من الصحة validation والاختبار test؟

ج/ بشكل عام، يعد وضع **80**٪ من البيانات في مجموعة التدريب و **10**٪ في مجموعة التحقق من الصحة و **10**٪ في مجموعة الاختبار تقسيمًا جيدًا للبدء به. يعتمد التقسيم الأمثل للاختبار والتحقق من الصحة ومجموعة التدريب على عوامل مثل حالة الاستخدام وهيكل النموذج وأبعاد البيانات وما إلى ذلك.

س84/ ما هي النسبة المثلى لبيانات التدريب والاختبار؟

ج/ النسبة شائعة الاستخدام هي **80:20**، مما يعني أن **80**٪ من البيانات مخصصة للتدريب و**20**٪ للاختبار. يتم أيضًا استخدام نسب أخرى مثل **70:30** و **60:40** وحتى **50:50** في الممارسة العملية.

س85/ ما هو استخدام ARIMA؟

ج/ **ARIMA** هو اختصار لعبارة "**autoregressive integrated moving average**". إنه نموذج يستخدم في الإحصاء والاقتصاد القياسي لقياس الأحداث التي تحدث خلال فترة زمنية. يستخدم النموذج لفهم البيانات السابقة أو التنبؤ بالبيانات المستقبلية في سلسلة.

س86/ هل ARIMA تعلماً عميقاً؟

ج/ نعم، هي واحدة من خوارزميات التعلم الآلي الأسهل والأكثر فعالية لأداء التنبؤ بالسلاسل الزمنية. هي مزيج من الانحدار التلقائي **Auto Regression** والمتوسط المتحرك **Moving average**.

س87/ ما هي مزايا نموذج ARIMA؟

ج/ الميزة الرئيسية لتنبؤ **ARIMA** هي أنها تتطلب بيانات عن السلاسل الزمنية المعنية فقط. أولاً، هذه الميزة مفيدة إذا كان المرء يتوقع عددًا كبيرًا من السلاسل الزمنية. ثانيًا، يتجنب هذا مشكلة تحدث أحيانًا مع النماذج متعددة المتغيرات **multivariate models**.

س88/ هل تستخدم Netflix التعلم العميق؟

ج/ تستخدم **Netflix** خوارزميات التعلم الآلي والعميق للتنبؤ بأنماط العارض **viewer patterns** وفهم متى سيكون هناك زيادات ونقصان عامة في عدد المشاهدين من الارتفاع المفاجئ في مشاهدة فيلم أو عرض معين.

س89/ هل يستخدم Siri التعلم الآلي؟

ج/ يعتمد **Siri** على توليد اللغة الطبيعية ومعالجة اللغة الطبيعية **NLP** والتعلم الآلي من أجل العمل بفعالية وتحسين أدائه بمرور الوقت.

س90/ ما هو استخدام VGG16؟

ج/ **VGG16** عبارة عن بنية شبكة عصبية التفافية **(CNN)** تم استخدامها للفوز بمسابقة **ILSVRC (Imagenet)** في عام **2014**.

س91/ ما هو الفرق بين CNN وVGG16؟

ج/ **CNN** هو مفهوم للشبكة العصبية، وقد تكون سماته الرئيسية أنه يتكون من طبقات الالتفاف، وطبقات التجميع، وطبقات التنشيط وما إلى ذلك. **VGG** هي شبكة تلافيفية **CNN** محددة مصممة للتصنيف **classification** والتوطين **localization**.

س92/ ما هي النماذج المدربة مسبقاً Pretrained model؟

ج/ النموذج المدرَّب مسبقًا **Pretrained model** عبارة عن شبكة محفوظة تم تدريبها مسبقًا على مجموعة بيانات كبيرة، عادةً في مهمة تصنيف صور على نطاق واسع. إما أن تستخدم النموذج الذي تم اختباره مسبقًا كما هو أو تستخدم نقل التعلم **transfer learning** لتخصيص هذا النموذج لمهمة معينة.

س93/ هل VGG16 نموذج مدرب مسبقًا Pretrained model؟

ج/ نعم, **VGG16** هو نموذج مدرب مسبقاً, يمكننا استخدام شبكته المدربة مسبقًا للقيام بمهمة مشابهة.

س94/ كم عدد المعلمات parameters هل VGG16؟

ج/ يحتوي **VGG16** على إجمالي **138** مليون معلمة. النقطة المهمة التي يجب ملاحظتها هنا هي أن جميع نواة الالتفاف بحجم **3 × 3** وأن نواة **maxpool** بحجم **2 × 2** بخطوة **stride** من اثنين.

س95/ كيف نستخدم VGG16 في التعرف على الوجوه face recognition؟

ج/ التعرف على الوجوه باستخدام نقل التعلم مع **VGG16**:

- الخطوة **1**: اجمع مجموعة البيانات **dataset**. لإنشاء أي نموذج ، فإن المطلب الأساسي هو مجموعة البيانات. لذلك دعونا نجمع بعض البيانات.
- الخطوة **2**: تدريب النموذج باستخدام **VGG16**. قم بتحميل أوزان **VGG16** وقم بتجميدها.
- الخطوة **3**: اختبر النموذج وقم بتشغيله. قم بتحميل النموذج لغرض الاختبار.

س96/ كيف يمكنني تحسين نموذج VGG16 الخاص بي؟

ج/ قم بتحسين شبكة **VGG16** باستخدام نواة التفاف **convolution kernel** كبيرة بدلاً من نواة التفاف صغيرة وتقليل بعض الطبقات المتصلة **fully connected layers** بالكامل لتقليل تعقيد النموذج ومعلماته.

س97/ كم عدد الخلايا العصبية في VGG؟

ج/ هناك أكثر من **60** مليون معلمة **parameters** و **650.000** خلية عصبية **neurons** تشارك في معمارية **VGG16**.

VGG16 إلى **16** طبقة لها أوزان. يوجد في **VGG16** ثلاثة عشر طبقة تلافيفية **convolutional layers** وخمس طبقات تجميع **Max Pooling layers** وثلاث طبقات كثيفة **Dense layers** تصل إلى **21** طبقة ولكنها تحتوي فقط على ستة عشر طبقة وزن أي طبقة معلمات قابلة للتعلم.

س98/ كيف يمكنني تحسين دقة VGG16 الخاصة بي؟

ج/ بالنسبة للأشخاص الذين قد يعانون من مشكلات في الدقة، يمكنك تجربة ما يلي:

- تحميل أوزان **VGG-16** المدربة مسبقًا.
- اجعل الطبقات التلافيفية القليلة الأخيرة قابلة للتدريب.
- استخدم مُحسِّن **SGD** وضبط معدل التعلم منخفضًا.
- ضبط دالة التنشيط الصحيحة في طبقة الإخراج.
- زيادة الفترات.

س99/ هل ResNet هي CNN؟

ج/ شبكة **ResNet** هي بنية شبكة عصبية تلافيفية (**CNN**) تغلبت على مشكلة "التدرج المتلاشي **vanishing gradient** "، مما يجعل من الممكن إنشاء شبكات بما يصل إلى آلاف الطبقات التلافيفية ، والتي تتفوق في الأداء على الشبكات الضحلة **shallow networks**.

س100/ كيف تختلف ResNet عن CNN؟

ج/ **ResNet** هي طريقة للتعامل مع مشكلة تلاشي التدرج في شبكات **CNN** العميقة جدًا. إنهم يعملون عن طريق تخطي بعض الطبقات بافتراض حقيقة أن الشبكات العميقة جدًا لا ينبغي أن تنتج خطأ تدريب أعلى من نظيراتها الضحلة. من منظور شامل، يمكن اعتبارها نموذجًا مشابهًا لـ **LSTM** في **RNNs**.

س101/ ما هي معمارية ResNet50؟

ج/ يتكون نموذج **ResNet-50** من **5** مراحل لكل منها كتلة التفاف **convolution** وهوية **Identity**. تحتوي كل كتلة التفاف على **3** طبقات التفاف ولكل كتلة الهوية أيضًا **3** طبقات التفاف. يحتوي **ResNet-50** على أكثر من **23** مليون معلمة قابلة للتدريب.

س102/ ما هو حجم الإدخال لـ ResNet50؟

ج/ حجم الادخال هو **224 x 224 x 3** .

س103/ ما هو استخدام ResNet-50؟

ج/ **ResNet-50** عبارة عن شبكة عصبية تلافيفية **CNN** بعمق **50** طبقة. يمكنك تحميل نسخة مسبقة التدريب من الشبكة تم تدريبها على أكثر من مليون صورة من قاعدة بيانات **ImageNet**. يمكن للشبكة التي تم اختبارها مسبقًا تصنيف الصور إلى **1000** فئة كائن، مثل لوحة المفاتيح والماوس والقلم الرصاص والعديد من الحيوانات.

س104/ هل ResNet50 هو نقل التعلم transfer learning؟

ج/ نعم، نحن نستخدم نموذج التعلم العميق **ResNet50** كنموذج مدرب مسبقًا لاستخراج الميزات من أجل نقل التعلم.

~~س105/ أيهما أفضل VGG16 أم ResNet؟~~

~~ج/ شبكة VGG-16 يحتوي على ما يقرب من 138 مليون معلمة وأن شبكة ResNet بها 25.5 مليون معلمة وبسبب هذا فهي أسرع، وهذا ليس صحيحًا. يقلل عدد المعلمات من مقدار المساحة المطلوبة لتخزين الشبكة، ولكن هذا لا يعني أنها أسرع.~~

س106/ لماذا تعتبر ResNet جيدة لتصنيف الصور؟

ج/ يتكون ResNet-50 من عمق 50 طبقة ويتم تدريبه على مليون صورة من 1000 فئة من قاعدة بيانات ImageNet. علاوة على ذلك، يحتوي النموذج على أكثر من 23 مليون معلمة قابلة للتدريب ، مما يشير إلى معمارية عميقة تجعله أفضل للتعرف على الصور.

~~س107/ لماذا تعتبر ResNet جيدة لتصنيف الصور؟~~

~~ج/ يتكون ResNet-50 من عمق 50 طبقة ويتم تدريبه على مليون صورة من 1000 فئة من قاعدة بيانات ImageNet. علاوة على ذلك، يحتوي النموذج على أكثر من 23 مليون معلمة قابلة للتدريب ، مما يشير إلى معمارية عميقة تجعله أفضل للتعرف على الصور.~~

س108/ ما هو نموذج Inception v3؟

ج/ Inception v3 هو نموذج مدرب مسبقاً للتعرف على الصور ثبت أنه يحقق دقة تزيد عن 78.1٪ على مجموعة بيانات ImageNet. النموذج هو تتويج للعديد من الأفكار التي طورها العديد من الباحثين على مر السنين.

س109/ هل Inception v3 هو نموذج CNN؟

ج/ Inception-v3 هو نموذج شبكة عصبية تلافيفية CNN مدرب مسبقًا وعمق 48 طبقة.

س110/ كم عدد الطبقات الموجودة في Inception-v3؟

ج/ Inception-v3 هي شبكة عصبية تلافيفية بعمق 48 طبقة. يمكنك تحميل نسخة مسبقة التدريب من الشبكة تم تدريبها على أكثر من مليون صورة من قاعدة بيانات ImageNet. يمكن للشبكة التي تم اختبارها مسبقًا تصنيف الصور إلى 1000 فئة كائن، مثل لوحة المفاتيح والماوس والقلم الرصاص والعديد من الحيوانات.

س111/ ما هو الفرق بين inception v2 و v3؟

ج/ يشبه **Inception V3** ويحتوي على جميع ميزات **Inception V2** مع التغييرات / الإضافات التالية: استخدام مُحسِّن **RMSprop**. تسوية الدُفعات في الطبقة المتصلة بالكامل من المصنف الإضافي. استخدام **7 × 7** فلتر التفاف.

س112/ هل Inception-v3 أفضل من VGG16؟

ج/ يحتوي **vgg16** على عدد من المعلمات أكبر من **inceptionv3** ، لذلك من الأسهل ملاءمة البيانات الجديدة إلى حد ما بسبب السعة الزائدة. شيء آخر نضيفه هو أن عمق **inception v3** أعمق بكثير من **VGG16**.

س113/ ما هو حجم الإدخال input size في inception-v3؟

ج/ حجم صورة الإدخال الافتراضي لـ **Inception-v3** هو **299 × 299**.

س114/ هل RNN خاضعة للأشراف supervised ام غير خاضعة للأشراف unsupervised؟

ج/ **RNN** هو نوع من التعلم العميق الخاضع للإشراف حيث يتم تغذية الإخراج من الخطوة السابقة كمدخل إلى الخطوة الحالية. خوارزمية التعلم العميق **RNN** هي الأنسب للبيانات المتسلسلة.

س115/ ما هي أنواع الـ RNN؟

ج/ أنواع الـ **RNN**:

- One-to-One RNN.
- One-to-Many RNN.
- Many-to-One RNN.
- Many-to-Many RNN.

س116/ ما هي عيوب الـ RNN؟

ج/ عيوب **RNN**:

- مشكلة تدريب **RNNs**.
- مشكلة التدرج المتلاشي أو المتفجر.
- لا يمكن تكديس **RNNs**.
- إجراءات تدريبية بطيئة ومعقدة.
- من الصعب معالجة التسلسلات الأطول.

س117/ هل يمكننا استخدام RNN لتصنيف الصور؟

ج/ .

س118/ ما نوع البيانات المناسبة لـ RNN؟

ج/ الشبكات العصبية المتكررة (**RNN**) هي فئة من الشبكات العصبية التي تساعد في نمذجة البيانات التسلسلية **sequence data**. تنتج الشبكات العصبية المتكررة نتائج تنبؤية في بيانات متسلسلة لا تستطيع الخوارزميات الأخرى تحقيقها.

س119/ كيف تختلف RNN عن الشبكات العصبية الأخرى؟

ج/ يتمثل الاختلاف الرئيسي بين **CNN** و**RNN** في القدرة على معالجة المعلومات الزمنية **temporal information** – البيانات التي تأتي في تسلسل، مثل الجملة. تم تصميم الشبكات العصبية المتكررة لهذا الغرض بالذات، في حين أن الشبكات العصبية التلافيفية **CNN** غير قادرة على تفسير المعلومات الزمنية بشكل فعال.

س120/ هل يمكننا الجمع بين RNN وCNN؟

ج/ نعم, يمكن الجمع بين **RNN** و **CNN**، والتي ستتيح الاستفادة الكاملة من المزايا الخاصة بكل منها: يمكن لـ **RNN** تعلم الميزات الزمنية والسياقية ، وخاصة التبعية طويلة المدى بين كيانين ، في حين أن **CNN** قادرة على التقاط المزيد من الميزات المحتملة.

س121/ هل CNN أسرع من RNN؟

ج/ تُعد شبكات **CNN** أسرع من شبكات **RNN** لأنها مصممة للتعامل مع الصور، بينما تم تصميم شبكات **RNN** للتعامل مع النص. بينما يمكن تدريب **RNNs** على التعامل مع الصور، لا يزال من الصعب عليهم فصل الميزات المتباينة القريبة من بعضها البعض

~~س122/ لماذا يتم استخدام الـ RNN للبيانات النصية؟~~

~~ج/ يتم تدريب **RNN** على التعرف على الأنماط عبر الزمن **patterns across time**، بينما تتعلم **CNN** التعرف على الأنماط عبر الفضاء **patterns across space**. يعتمد نوع **DNN** الذي يعمل بشكل أفضل عند التعامل مع البيانات النصية على عدد المرات المطلوبة لفهم الدلالات العالمية/بعيدة المدى.~~

س123/ هل يمكننا استخدام الشبكة العصبية المتكررة (RNN) للتصنيف الثنائي؟

ج/ نعم، ولكن عادةً ما تعمل **RNN** بشكل أفضل مع بيانات السلاسل الزمنية حيث يلزم دمج المعلومات السابقة. ولكن إذا كان التصنيف الوحيد هو الهدف النهائي وكانت البيانات عبارة عن سلاسل غير زمنية، فيجب أن تكون خوارزمية بسيطة من الانحدار اللوجستي للتصنيف الثنائي كافية لأنها ستقلل من تعقيد خوارزمية التنفيذ.

س124/ لماذا RNN أفضل من CNN لـ NLP؟

ج/يتم تدريب **RNN** على التعرف على الأنماط عبر الزمن **patterns across time**، بينما تتعلم **CNN** التعرف على الأنماط عبر الفضاء **patterns across space**. يعتمد نوع **DNN** الذي يعمل بشكل أفضل عند التعامل مع البيانات النصية على عدد المرات المطلوبة لفهم الدلالات العالمية/بعيدة المدى.

س125/ ما هي المشاكل التي يمكن لـ RNN أفضل من الـ CNN؟

ج/ يمكننا استخدام الشبكات العصبية المتكررة **RNN** لحل المشكلات المتعلقة بـ: بيانات السلاسل الزمنية. بيانات نصية. البيانات الصوتية.

س126/ ما هو ال RNN في معالجة الصور؟

ج/ من غير المحتمل أن تجد العديد من الأمثلة على استخدام **RNN** لتصنيف صورة لأن **RNNs** أدنى من شبكات **CNN** في معظم مهام معالجة الصور. قد تستخدم **RNN** للأغراض التعليمية أكثر من الأغراض العملية.

س127/ كيف يتم تدريب RNN؟

ج/ لتدريب **RNNs**، يتم استخدام الانتشار الخلفي عبر الزمن **Backpropagation Through Time (BPTT)**. يتم إلحاق كلمة "عبر الزمن" بمصطلح "**backpropagation**" لتحديد أن الخوارزمية يتم تطبيقها على نموذج عصبي زمني (**RNN**). تتمثل مهمة **BPTT** في إيجاد حد أدنى محلي، نقطة بها أقل خطأ. من خلال ضبط قيم الأوزان، يمكن أن تصل الشبكة إلى الحد الأدنى.

س128/ أيهما أفضل لتصنيف الصور CNN أم RNN؟

ج/ تعمل شبكات **CNN** بشكل أفضل مع البيانات المكانية **spatial data** ، وبالتالي فهي الخيار الأنسب لمعالجة الصور والفيديو. من ناحية أخرى، تعمل **RNN** على بيانات متسلسلة **sequential data** وبالتالي تثبت أنها خيار مناسب لتحليل النص والكلام.

س129/ ما هي الشبكة العصبية العميقة الأفضل لـ chatbot؟

ج/ **RNN** هي واحدة من أكثر أنواع الشبكات العصبية استخدامًا، ويرجع ذلك أساسًا إلى قدرتها الكبيرة على التعلم وقدرتها على أداء المهام المعقدة مثل تعلم الكتابة اليدوية أو التعرف على اللغة.

س130/ كيف يتم استخدام RNN في chatbot؟

ج/ يتصور المشفر لـ **RNN** سلسلة من الرموز المميزة للسياق واحدًا تلو الآخر ويقوم بتحديث حالته المخفية. بعد معالجة تسلسل السياق بالكامل، ينتج عنه حالة خفية نهائية، والتي تتضمن الإحساس بالسياق وتستخدم لتوليد الإجابة.

س131/ ما هو LSTM ولماذا يتم استخدامه؟

ج/ الذاكرة طويلة قصيرة المدى (**LSTM**) هي شبكة عصبية اصطناعية تستخدم في مجالات الذكاء الاصطناعي والتعلم العميق. على عكس الشبكات العصبية ذات التغذية الأمامية القياسية، فإن **LSTM** لديها اتصالات تغذية مرتدة **has feedback connections**. تعتبر شبكات **LSTM** مناسبة تمامًا لتصنيف ومعالجة وعمل التنبؤات بناءً على بيانات السلاسل الزمنية، حيث يمكن أن يكون هناك فترات تأخير غير معروفة بين الأحداث المهمة في سلسلة زمنية. تم تطوير **LSTMs** للتعامل مع مشكلة التدرج المتلاشي **vanishing gradient** التي يمكن مواجهتها عند تدريب **RNNs** التقليدية.

س132/ لماذا LSTM أفضل من RNN؟

ج/ شبكات **LSTM** هي نوع من شبكات **RNN** التي تستخدم وحدات خاصة بالإضافة إلى الوحدات القياسية. تتضمن وحدات **LSTM** "خلية ذاكرة **memory cell** " يمكنها الاحتفاظ بالمعلومات في الذاكرة لفترات طويلة من الزمن. تتيح لهم خلية الذاكرة هذه تعلم التبعيات طويلة المدى **longer-term dependencies**.

س133/ كيف تعمل LSTM؟

ج/ تستخدم **LSTMs** سلسلة من "البوابات **gates**" التي تتحكم في كيفية وصول المعلومات الموجودة في سلسلة من البيانات إلى الشبكة وتخزينها فيها ومغادرتها. توجد ثلاث بوابات في **LSTM** النموذجي؛ بوابة النسيان **forget gate** وبوابة الإدخال **input gate** وبوابة الإخراج **output gate**. يمكن اعتبار هذه البوابات بمثابة مرشحات **filters** وكل منها شبكة عصبية خاصة بها.

س134/ لماذا LSTM هو الأفضل للسلسلة الزمنية time series؟

ج/ باستخدام **LSTM**، يمكن لنماذج التنبؤ بالسلاسل الزمنية التنبؤ بالقيم المستقبلية بناءً على البيانات المتسلسلة السابقة. يوفر هذا دقة أكبر للمتنبئين بالطلب **demand forecasters** مما يؤدي إلى اتخاذ قرارات أفضل للأعمال.

س135/ كم عدد الطبقات التي تمتلكها LSTM؟

ج/ تتكون شبكة **LSTM** من ثلاث طبقات؛ طبقة إدخال، طبقة مخفية واحدة متبوعة بطبقة إخراج أمامية قياسية.

س136/ ما المشكلة التي تحلها LSTM؟

ج/ تم ايجاد **LSTM** خصيصًا لتجنب مشكلة التدرج المتلاشي **vanishing gradient** التي يمكن مواجهتها عند تدريب **RNNs** التقليدية.

س137/ لماذا يتم استخدام LSTM في RNN؟

ج/ تُستخدم وحدات **LSTM** كوحدات بناء لطبقات **RNN**، وغالبًا ما تسمى شبكة **LSTM**. تمكن **LSTMs** الـ **RNN** من تذكر المدخلات على مدى فترة طويلة من الزمن. هذا لأن **LSTMs** تحتوي على معلومات في الذاكرة، مثل ذاكرة الكمبيوتر.

س138/ ما هي الاكثر العيوب التي تعاني منها الـ LSTM؟

ج/ أحد القيود المهمة على **LSTM** هو الذاكرة. أو بشكل أكثر دقة، كيف يمكن إساءة استخدام الذاكرة. باختصار، تتطلب **LSTM 4** طبقات خطية (طبقة **MLP**) لكل خلية للتشغيل عند ولكل خطوة زمنية للتسلسل. تتطلب الطبقات الخطية كميات كبيرة من عرض النطاق الترددي للذاكرة ليتم حسابها، وفي الواقع لا يمكنها استخدام العديد من وحدات الحوسبة في كثير من الأحيان لأن النظام ليس لديه عرض نطاق ذاكرة كافٍ لتغذية الوحدات الحسابية

س139/ ما هي عيوب LSTM؟

ج/ **LSTM** لديها بعض العيوب منها:

- تستغرق **LSTMs** وقتًا أطول للتدريب.
- تتطلب **LSTMs** مزيدًا من الذاكرة للتدريب.
- **LSTMs** عرضة للضبط الزائد **overfitting**.
- **Dropout** أصعب بكثير في التنفيذ في **LSTM**.
- **LSTMs** حساسة لمختلف عمليات تهيئة الوزن العشوائية.

س140/ ما هي المعلمات parameters في LSTM؟

ج/ تحتوي طبقة **LSTM** على "أبعاد مساحة الإخراج **imensionality of the output space**" (وحدة) قيمة المعلمة **2** مما يعني أن الحالات المخفية والخلية عبارة عن متجهات ذات البعد **2**. يتم تمثيل الإدخال لكل خطوة زمنية بواسطة متجه مع البعد **3** (ميزة).

س141/ ما هو الفرق بين CNN وLSTM؟

ج/ تم تصميم **LSTM** للعمل بشكل مختلف عن **CNN** لأنه يتم استخدام **LSTM** عادةً لمعالجة وتكوين تنبؤات معطاة تسلسل البيانات (على النقيض من ذلك ، تم تصميم **CNN** لاستغلال "الارتباط المكاني **spatial correlation** " في البيانات ويعمل بشكل جيد على الصور والكلام).

س142/ هل يمكننا تطبيق نموذج LSTM لتصنيف الصور؟

ج/ نعم، يمكن تطبيق نموذج **LSTM** لتصنيف الصور. لكن يفضل جمعها مع الـ **CNN** لتصنيف الصور.

س143/ هل LSTM جيدة للتصنيف؟

ج/ بشكل عام، يمكن استخدام **LSTM** للتصنيف **classification** أو الانحدار **regression**؛ إنها في الأساس مجرد شبكة عصبية **NN** قياسية تأخذ كمدخلات، بالإضافة إلى المدخلات من تلك الخطوة الزمنية **time step**، حالة مخفية من الخطوة الزمنية السابقة. لذلك، مثلما يمكن استخدام الشبكات العصبية **NN** للتصنيف أو الانحدار، كذلك يمكن استخدام **LSTM**.

س144/ كيف يتجنب LSTM مشكلة تلاشي الانحدار؟

ج/ تحل **LSTMs** المشكلة باستخدام بنية انحدار **gradient** مضافة فريدة تتضمن الوصول المباشر إلى عمليات تفعيل بوابة النسيان **forget gate**، مما يمكّن الشبكة من تشجيع السلوك المطلوب من خطأ الانحدار باستخدام تحديث البوابات المتكرر في كل خطوة زمنية من عملية التعلم.

س145/ ما هي مدخلات ومخرجات LSTM؟

ج/ دائمًا ما يكون إدخال **LSTM** عبارة عن مصفوفة ثلاثية الأبعاد. (حجم_الدفعة ، الخطوات_الزمنية ، **seq_len**). يمكن أن يكون ناتج **LSTM** عبارة عن مصفوفة ثنائي الأبعاد أو مصفوفة ثلاثية الأبعاد اعتمادًا على وسيطة **return_sequences**.

س146/ كيف تزيد الدقة accurecy في LSTM؟

ج/ هناك خياران لزيادة الدقة:

1) زيادة الطبقات المخفية في عقدة **LSTM**.
2) إضافة طبقة أخرى من **LSTM**.

س147/ لماذا يتم استخدام LSTM في البرمجة اللغوية العصبية NLP؟

ج/ سهلت لنا **LSTM** إعطاء جملة كمدخل للتنبؤ بدلاً من مجرد كلمة واحدة، وهو أكثر ملاءمة في البرمجة اللغوية العصبية ويجعلها أكثر كفاءة.

س148/ كيف يعمل LSTM خطوة بخطوة؟

ج/ تحتوي مصفوفة الوزن **W** على أوزان مختلفة لمتجه الإدخال الحالي والحالة المخفية السابقة لكل بوابة. تمامًا مثل الشبكات العصبية المتكررة **RNN**، تولد شبكة **LSTM** أيضًا ناتجًا في كل خطوة زمنية ويتم استخدام هذا الإخراج لتدريب الشبكة باستخدام الانحدار الاشتقاقي **gradient descent**.

س149/ كيف يتم تدريب LSTM؟

ج/ يتم تدريب **LSTM** (ضبط المعلمات) مع نافذة إدخال للبيانات السابقة وتقليل الفرق بين القيمة المتوقعة والقيمة المقاسة التالية. تتنبأ الطرق المتسلسلة بقيمة تالية واحدة فقط بناءً على نافذة البيانات السابقة.

س150/ ما هو استخدام GRU؟

ج/ الوحدة المتكررة ذات البوابات **Gated Recurrent Unit (GRU)** هي جزء من نموذج محدد للشبكة العصبية المتكررة **RNN** التي تهدف إلى استخدام الاتصالات من خلال سلسلة من العقد لأداء مهام التعلم الآلي المرتبطة بالذاكرة والتكتل **clustering**، على سبيل المثال، في التعرف على الكلام **speech recognition**.

س151/ كيف تختلف GRU عن RNN القياسية؟

ج/ سير عمل الوحدة المتكررة ذات البوابات **Gated Recurrent Unit**، باختصار **GRU**، هو نفسه **RNN** ولكن الاختلاف في العملية والبوابات المرتبطة بكل وحدة **GRU**. لحل المشكلة التي تواجهها **RNN** القياسية، تدمج **GRU** آليتي تشغيل البوابة المسماة بوابة التحديث **Update gate** وبوابة إعادة التعيين **Reset gate**.

س152/ أيهما أفضل LSTM أو GRU؟

ج/ يتمثل الاختلاف الرئيسي بين **GRU** و**LSTM** في أن حقيبة **GRU** بها بوابتين يتم إعادة تعيينهما **Reset** وتحديثهما **Update** بينما تحتوي **LSTM** على ثلاث بوابات هي الإدخال والإخراج والنسيان. **GRU** أقل تعقيدًا من **LSTM** لأنه يحتوي على عدد أقل من البوابات. إذا كانت مجموعة البيانات صغيرة، فيُفضل استخدام **GRU**، وإلا فإن **LSTM** لمجموعة البيانات الأكبر. تستخدم **GRU** معلمة تدريب أقل وبالتالي تستخدم ذاكرة أقل ويتم تنفيذها بشكل أسرع من **LSTM** بينما يكون **LSTM** أكثر دقة في مجموعة بيانات أكبر.

س153/ كيف يمكنني إيقاف الضبط الزائد overfitting في LSTM؟

ج/ قلل عدد الوحدات في **LSTM** حتى تصل إلى نقطة يتوقف عندها نموذجك عن الضبط الزائد. ثم أضف **dropout** إذا لزم الأمر.

س154/ ما هو نموذج LSTM ثنائي الاتجاه Bi-LSTM؟

ج/ الذاكرة طويلة قصيرة المدى ثنائية الاتجاه (**Bi-LSTM**) هي عملية جعل أي شبكة عصبية لديها معلومات التسلسل في كلا الاتجاهين للخلف (من المستقبل إلى الماضي) أو للأمام (من الماضي إلى المستقبل). في الاتجاهين، تتدفق مدخلاتنا في اتجاهين ، مما يجعل **Bi-LSTM** مختلفًا عن **LSTM** العادي.

س155/ ما هو الفرق بين LSTM و Bi-LSTM؟

ج/ سيؤدي استخدام **Bi-LSTM** إلى تشغيل مدخلاتك بطريقتين ، واحدة من الماضي إلى المستقبل والأخرى من المستقبل إلى الماضي وما يختلف هذا النهج من الاتجاه الأحادي هو أنه في **LSTM** الذي يعمل بشكل عكسي ، فإنك تحتفظ بالمعلومات من المستقبل واستخدام الحالتين المخفيتين معًا قادرون في أي وقت على الاحتفاظ بالمعلومات من الماضي والمستقبل.

س156/ أيهما أفضل LSTM أم BiLSTM؟

ج/ تظهر النتائج أن التدريب الإضافي للبيانات وبالتالي النمذجة المستندة إلى **BiLSTM** تقدم تنبؤات أفضل من النماذج العادية القائمة على **LSTM**. وبشكل أكثر تحديدًا، لوحظ أن نماذج **BiLSTM** توفر تنبؤات أفضل مقارنة بنماذج **LSTM**.

س157/ ما هو استخدام LSTM ثنائي الاتجاه؟

ج/ عادةً ما يتم استخدام **BI-LSTM** عند الحاجة إلى تسلسل المهام **sequence tasks**. يمكن استخدام هذا النوع من الشبكات في تصنيف النصوص **text classification** والتعرف على الكلام **speech recognition** ونماذج التنبؤ **forecasting models**.

س158/ لماذا LSTM بطيئة؟

ج/ هذا يرجع بشكل أساسي إلى الحساب المتسلسل في طبقة **LSTM**. تذكر أن **LSTM** يتطلب إدخالًا تسلسليًا لحساب أوزان الطبقة المخفية بشكل تكراري ، بمعنى آخر ، يجب أن تنتظر الحالة المخفية في الوقت **t-1** لحساب الحالة المخفية في الوقت **t**.

س159/ هل تحتاج إلى GPU لـ LSTM؟

ج/ تعد وحدات معالجة الرسوميات **GPU** هي المعيار الفعلي لاستخدام **LSTM** وتوفر تسريعًا بمعدل **6** أضعاف أثناء التدريب وإنتاجية أعلى بمقدار **140** ضعفًا أثناء الاستدلال عند مقارنتها بتطبيقات وحدة المعالجة المركزية. **cuDNN** عبارة عن مكتبة شبكة عصبية عميقة يتم تسريعها بواسطة **GPU** تدعم تدريب الشبكات العصبية المتكررة **LSTM** للتعلم المتسلسل.

س160/ كم عدد طبقات LSTM التي أحتاجها؟

ج/ بشكل عام، أظهرت طبقتان أنهما كافيتان لاكتشاف ميزات أكثر تعقيدًا. يمكن أن يكون تدريب المزيد من الطبقات أفضل ولكن يصعب تدريبه أيضًا. كقاعدة عامة، تعمل طبقة واحدة مخفية مع مشاكل بسيطة، مثل هذه، واثنتان تكفيان للعثور على ميزات معقدة بشكل معقول.

س161/ ماهي vanilla LSTM؟

ج/ **Vanilla LSTM** هو نموذج **LSTM** يحتوي على طبقة مخفية واحدة من وحدات **LSTM**، وطبقة مخرجات تستخدم لعمل تنبؤ. يمكننا تحديد **Vanilla LSTM** للتنبؤ بالسلسلة الزمنية أحادية المتغير على النحو التالي.

س162/ لماذا LSTM أفضل من Arima؟

ج/ يعمل **LSTM** بشكل أفضل إذا كنا نتعامل مع كمية هائلة من البيانات وتتوفر بيانات تدريب كافية، في حين أن **ARIMA** أفضل لمجموعات البيانات الأصغر تتطلب **ARIMA** سلسلة من المعلمات (**p**، **q**، **d**) والتي يجب حسابها على أساس على البيانات، بينما لا يتطلب **LSTM** تحديد مثل هذه المعلمات.

س163/ هل تدريب GRU أسهل من تدريب LSTM التقليدي؟

ج/ يعتبر تدريب **GRUs** اسهل لأنها تحتوي على معلمات أقل قابلية للضبط. في **GRU**، يتم دمج خلايا الإخراج والذاكرة. يتم دمج بوابتي "نسيان" و "إدخال" في بوابة "تحديث" واحدة.

س164/ هل تمتلك GRU وحدة ذاكرة؟

ج/ نعم تمتلك، تشترك **GRU** في العديد من خصائص الذاكرة طويلة قصيرة المدى **(LSTM)**. تستخدم كلتا الخوارزميتين آلية بوابة للتحكم في عملية الحفظ **memorization process**.

س165/ كم عدد المعلمات parameters في GRU؟

ج/ العدد الإجمالي للمعلمات في **GRU** يساوي **3 × (n2 + nm + n)**. حيث **m** هو بُعد الإدخال و **n** هو بُعد الإخراج.

س166/ كم عدد بوابات GRU؟

ج/ هناك بوابتان في **GRU** بدلاً من ثلاث بوابات في خلية **LSTM**. البوابة الأولى هي بوابة إعادة الضبط **Reset gate** والأخرى هي بوابة التحديث **update gate**.

س167/ ما هو المحول transformer في الشبكة العصبية؟

ج/ المحول **transformer** هو أحد المكونات المستخدمة في العديد من تصميمات الشبكات العصبية لمعالجة البيانات المتسلسلة، مثل نص اللغة الطبيعية أو تسلسل الجينوم أو الإشارات الصوتية أو بيانات السلاسل الزمنية. تقع معظم تطبيقات الشبكات العصبية للمحولات في مجال معالجة اللغة الطبيعية.

س168/ ماهي استخدامات المحول transformer؟

ج/ المحول هو نموذج التعلم العميق الذي يعتمد آلية الانتباه الذاتي **self-attention**، يهيئ بشكل تفاضلي اوزان لأهمية كل جزء من بيانات الإدخال. يتم استخدامه بشكل أساسي في مجالات معالجة اللغة الطبيعية **(NLP)** ورؤية الكمبيوتر **(CV)**.

س169/ ما هي المحولات في البرمجة اللغوية العصبية NLP؟

ج/ المحول هو أحد المكونات المستخدمة في العديد من تصميمات الشبكات العصبية لمعالجة البيانات المتسلسلة، مثل نص اللغة الطبيعية أو تسلسل الجينوم أو الإشارات الصوتية أو بيانات السلاسل الزمنية. تقع معظم تطبيقات الشبكات العصبية للمحولات في مجال معالجة اللغة الطبيعية **NLP**.

س170/ هل يمكن استخدام المحولات لتصنيف النص text classification؟

ج/ نموذج المحول قادر على الأداء بشكل جيد في مهمة تصنيف النص حيث أننا قادرون على تحقيق النتائج المرجوة في معظم تنبؤاتنا.

س171/ ما هي خوارزمية BERT؟

ج/ خوارزمية **BERT** (تمثيلات التشفير ثنائية الاتجاه من المحولات) هي خوارزمية التعلم العميق المتعلقة بمعالجة اللغة الطبيعية. إنه يساعد الآلة على فهم معنى الكلمات في الجملة، ولكن مع كل الفروق الدقيقة في السياق.

س172/ ما هي تطبيقات BERT؟

ج/يمكن استخدام **BERT** في مجموعة متنوعة من مهام البرمجة اللغوية العصبية **NLP** مثل تصنيف النص **Text Classification** أو تصنيف الجمل **Sentence Classification**، والتشابه الدلالي بين أزواج الجمل **Semantic Similarity between pairs of Sentences**، ومهمة الإجابة على الأسئلة مع الفقرة ، وتلخيص النص **text summarization** ، إلخ.

س173/ هل يمكن استخدام BERT للترجمة؟

ج/ **BERT** ليس نموذجًا للترجمة الآلية **machine translation**، لقد تم تصميم **BERT** لتوفير تمثيل الجملة السياقية الذي يجب أن يكون مفيدًا لمختلف مهام البرمجة اللغوية العصبية.

س174/ هل يمكن أن تحل المحولات محل LSTM؟

ج/تم تقديم المحولات في عام **2017** من قبل فريق في **Google Brain** وهي بشكل متزايد النموذج المفضل لمشاكل البرمجة اللغوية العصبية **NLP**، لتحل محل نماذج **RNN** مثل الذاكرة طويلة قصيرة المدى (**LSTM**). يسمح موازاة التدريب **training parallelization** الإضافي بالتدريب على مجموعات بيانات أكبر.

س175/ لماذا المحولات transformers أفضل من LSTM؟

ج/ نظرًا لان تصميم المحول **transformer** يمكن أن يسمح لكل من البيانات والتدريب المتوازي للنموذج، يكون المحول أكثر كفاءة من الشبكة العصبية المتكررة مثل **LSTM**. في الوقت نفسه، تم اقتراح بنية وحدة فك التشفير أيضًا لموازنة التأثير والكفاءة.

س176/ هل يعتمد BERT على المحولات؟

ج/ تستند **BERT**، التي تمثل تمثيلات التشفير ثنائية الاتجاه من **Transformers** ، إلى **Transformers** ، وهو نموذج تعلم عميق يرتبط فيه كل عنصر مخرجات بكل عنصر إدخال ، ويتم حساب الترجيحات بينها ديناميكيًا بناءً على اتصالها.

س177/ ما هو استخدام GAN؟

ج/ شبكة الخصومة التوليدية (GAN) هي نموذج للتعلم الآلي (ML) تتنافس فيه شبكتان عصبيتان مع بعضهما البعض لتصبحا أكثر دقة في تنبؤاتهما.

س178/ هل GAN لعبة مجموع صفري zero sum game؟

ج/ نظر إلى شبكات **GAN** عمومًا على أنها لعبة لاعبين محصلتها صفر **zero- sum game** بين شبكتين عصبيتين.

س179/ ما هي الأنواع المختلفة لشبكات GAN؟

ج/

1) DC GAN
2) CGAN
3) LSGAN
4) ACGAN
5) DVD-GAN
6) SRGAN
7) Cycle GAN
8) Info GAN

س180/ ما هي التحديات التي تواجهها شبكات الخصومة التوليدية (GANs)؟

ج/

1. مشكلة الثبات بين المولد والمميز: لا نريد أن يكون هذا التمييز صارمًا للغاية ، نريد أن نكون متساهلين
2. مشكلة تحديد مواقع الأشياء: افترض في الصورة أن لدينا **3** أحصنة وأن المولد أنشأ **6** عيون وحصان واحد.
3. مشكلة فهم الأشياء العالمية : لا تفهم شبكات **GAN** الهيكل العالمي أو الهيكل الشامل الذي يشبه مشكلة المنظور. هذا يعني في بعض الأحيان أن **GAN**تنشئ صورة غير واقعية ولا يمكن أن تكون ممكنة.
4. مشكلة في فهم المنظور : لا يمكن فهم الصور ثلاثية الأبعاد وإذا قمنا بتدريبها على مثل هذه الأنواع من الصور ، فسوف تفشل في إنشاء صور ثلاثية الأبعاد لأن شبكات **GAN** اليوم قادرة على العمل على صور ثلاثية الأبعاد.

س181/ ما هي خطوات تنفيذ شبكة GAN؟

ج/

1. استيراد جميع المكتبات
2. الحصول على مجموعة البيانات
3. إعداد البيانات - يتضمن خطوات مختلفة لإنجازها مثل المعالجة المسبقة للبيانات ، وتوسيع نطاق البيانات ، وتسويتها ، وإعادة تشكيلها.
4. تحديد وظيفة المولد والمميز.
5. قم بإنشاء ضوضاء عشوائية ثم قم بإنشاء صورة بها ضوضاء عشوائية.
6. إعداد المعلمات مثل تحديد الفترة ، وحجم الدفعة ، وحجم العينة.
7. تحديد وظيفة توليد صور العينة.
8. قم بتدريب المميز **Discriminator** ثم يقوم بتدريب المولد **Generator** وسوف يقوم بإنشاء الصور.
9. سوف نرى ما يتم إنشاؤه من وضوح الصور بواسطة المولد.

س182/ هل يمكن استخدام GAN للتصنيف؟

ج/ تم تطبيق شبكات **GAN** مؤخرًا على مهام التصنيف، وغالبًا ما تشترك في بنية واحدة لكل من التصنيف والتمييز. ومع ذلك، قد يتطلب هذا أن يتقارب النموذج مع توزيع منفصل للبيانات لكل مهمة ، مما قد يقلل من الأداء العام.

س183/ أي GAN هو الأفضل لتوليد الصور؟

ج/خمس شبكات **GAN** لتحسين معالجة الصور

- **Conditional GAN.**
- **Stacked GAN.**
- **Information Maximizing GAN.**
- **Super Resolution GAN.**
- **Pix2Pix.**

س184/ هل يمكن استخدام GAN للانحدار regression؟

ج/الجواب نعم. هناك ورقة بحثية بعنوان عزل التباين داخل الصنف في شبكات **GAN** المشروطة (**https://arxiv.org/pdf/1811.11296.pdf**) تناقشها. يقترحون "**C-GAN** قادرة على تعلم نماذج واقعية مع معلمات إدخال مستمرة وذات مغزى"

س185/ ما هو الفرق بين GAN وGAN الشرطي (C-GAN)؟

ج/في **GAN**، لا يوجد تحكم في أوضاع البيانات المراد إنشاؤها. يغير **GAN** الشرطي ذلك عن طريق إضافة التسمية **y** كمعامل إضافي للمولد ويأمل أن يتم إنشاء الصور المقابلة.

س186/ ما هو GAN ثنائي الاتجاه bidirectional GAN؟

ج/ **BiGAN** ، أو **GAN** ثنائية الاتجاه ، هي نوع من شبكات الخصومة التوليدية حيث لا يقوم المولد فقط بتعيين العينات الكامنة للبيانات التي تم إنشاؤها، ولكن لديه أيضًا تعيين عكسي من البيانات إلى التمثيل الكامن

س187/ لماذا يصعب تدريب GAN؟

ج/يصعب تدريب شبكات **GAN**. سبب صعوبة تدريبهم هو أن كلا من نموذج المولد ونموذج المميز يتم تدريبهما في وقت واحد في اللعبة. هذا يعني أن التحسينات على أحد النماذج تأتي على حساب النموذج الآخر.

س188/ هل GANهي التزييف العميق Deepfake؟

ج/تعتمد تقنية التزييف العميق عادةً على شبكات الخصومة التوليدية (**GANs**)، حيث يتم تدريب شبكتين عصبيتين متنافستين بشكل مشترك. حققت شبكات **GAN** نجاحًا كبيرًا في العديد من مهام رؤية الكمبيوتر.

س189/ ما هي عيوب GAN؟

ج/

- عدم التقارب **Non-convergence**: تتأرجح معلمات النموذج وتزعزع الاستقرار ولا تتقارب أبدًا،
- انهيار الوضع **Mode collapse**: ينهار المولد وينتج أنواعًا محدودة من العينات،
- التدرج المتناقص **Diminished gradient**: يصبح أداة المميز ناجحًا جدًا بحيث يتلاشى الانحدار التدريجي للمولد ولا يتعلم شيئًا،

س190/ كيف تزيد دقة GAN؟

ج/

- تغيير دالة التكلفة لتحقيق هدف تحسين أفضل.
- إضافة عقوبات إضافية إلى دالة التكلفة لفرض القيود.
- تجنب الضبط الناقص والزائد.

- طرق أفضل لتحسين النموذج.
- إضافة تسميات.

س191/ لماذا يستخدم Tanh في المولدات في GAN؟

ج/ هذا يرجع إلى حقيقة أنه عند إنشاء الصور، عادة ما يتم تطبيعها لتكون إما في النطاق [0،1] أو [-1،1]. لذا إذا كنت تريد أن تكون الصور الناتجة في [0،1] يمكنك استخدام **sigmoid** وإذا كنت تريد أن تكون في [-1،1] يمكنك استخدام **tanh**.

س192/ ما هو المحسن Optimizer الأفضل لـ GAN؟

ج/ محسن آدم **Adam optimizer**.

س193/ ماذا يعني أن خسارة أداة التمييز بقيت عند قيمة ثابتة بينما تتناقص خسارة المولد؟

ج/ يعني أن المولد أصبح أقوى من المميز. لتحقيق التوازن بين المولد والمميز، يمكنك محاولة تدريب المميز أكثر من المولد.

س194/ ما هو استخدام المشفر التلقائي autoencoder؟

ج/ تُستخدم المشفر التلقائي للمساعدة في تقليل الضوضاء في البيانات. من خلال عملية ضغط بيانات الإدخال وتشفيرها ثم إعادة بنائها كمخرج، تتيح لك المشفرات التلقائية تقليل الأبعاد والتركيز فقط على المجالات ذات القيمة الحقيقية.

س195/ هل يقلل المشفر التلقائي الأبعاد؟

ج/ تعمل المشفرات التلقائية كنماذج يمكنها تقليل أبعاد مدخلاتها. يتم تدريبهم على "محاكاة" مدخلاتهم. ينتج المشفر تمثيلًا كامنًا للمدخلات التي يتم تمريرها إلى وحدة فك التشفير، حيث يتم إعادة بناء المدخلات.

س196/ هل المشفر التلقائي غير خطي nonlinear؟

ج/ يقتصر **PCA** على خريطة خطية، بينما يمكن أن تحتوي المشفرات التلقائية على جهاز تشفير/مفكك تشفير غير خطي.

س197/ هل المشفر التلقائي خاضع للإشراف أم غير خاضع للإشراف؟

ج/ المشفر التلقائي هو نموذج شبكة عصبية يسعى إلى تعلم تمثيل مضغوط لمدخلات. إنها طريقة تعلم غير خاضعة للإشراف، على الرغم من أنها من الناحية الفنية ، يتم تدريبها باستخدام طرق التعلم الخاضعة للإشراف ، والتي يشار إليها باسم الإشراف الذاتي **self-supervised**.

س198/ ما هي الأنواع المختلفة للمشفر التلقائي؟

ج/ هناك، في الأساس، **7** أنواع من المشفرات التلقائية:

1. Denoising autoencoder.
2. Sparse Autoencoder.
3. Deep Autoencoder.
4. Contractive Autoencoder.
5. Undercomplete Autoencoder.
6. Convolutional Autoencoder.
7. Variational Autoencoder.

س199/ ما هو المشفر encoder في المشفر التلقائي Autoencoder؟

ج/ يتكون المشفر التلقائي من **3** مكونات: وحدة التشفير **encoder** والرمز **code** ووحدة فك التشفير **decoder**. يضغط المشفر المدخلات وينتج الكود، ثم يعيد مفكك التشفير بناء المدخلات فقط باستخدام هذا الكود.

س200/ كيف تقلل المشفرات التلقائية الخطأ؟

ج/ .

س201/ ما هو عنق الزجاجة bottleneck في التشفير التلقائي؟

ج/ عنق الزجاجة **bottleneck**: هي الطبقة المخفية ذات الأبعاد المنخفضة حيث يتم إنتاج الترميز. تحتوي طبقة عنق الزجاجة على عدد أقل من العقد كما أن عدد العقد في طبقة عنق الزجاجة يعطي أيضًا بُعد تشفير الإدخال.

س202/ ما هو الفضاء الكامنة latent space في المشفر التلقائي؟

ج/ الفضاء الكامن **latent space** هو مجرد تمثيل للبيانات المضغوطة حيث تكون نقاط البيانات المتشابهة أقرب معًا في الفضاء. يعد الفضاء الكامن مفيد لتعلم ميزات البيانات ولإيجاد تمثيلات أبسط للبيانات لتحليلها.

س203/ ما هي المشفرات التلقائية المتغيرة variational Autoencoder (VAE)؟

ج/ يمكن تعريف المشفر التلقائي **VAE** المتغير على أنه مشفر تلقائي يتم تنظيم تدريبه لتجنب الضبط الزائد **overfitting** والتأكد من أن المساحة الكامنة **latent space** لها خصائص جيدة تمكن من عملية التوليد.

س204/ ما هو الفرق بين المشفر التلقائي Autoencoder (AE) والمشفر التلقائي المتغير variational Autoencoder (VAE)؟

ج/ يخرج المشفر في AE النواقل الكامنة **latent vectors**. بدلاً من إخراج المتجهات في الفضاء الكامن **latent space**، يقوم مشفر VAE بإخراج معلمات توزيع محدد مسبقًا في المساحة الكامنة لكل إدخال. ثم يفرض VAE قيدًا على هذا التوزيع الكامن مما يجبره على أن يكون توزيعًا طبيعيًا..

س205/ لماذا نحتاج VAE؟

ج/ تتمثل الفائدة الرئيسية للمشفر التلقائي المتغير VAE في قدرتنا على تعلم تمثيلات الحالة الكامنة السلسة لبيانات الإدخال. بالنسبة للمشفرات التلقائية القياسية AE، نحتاج ببساطة إلى تعلم الترميز الذي يسمح لنا بإعادة إنتاج المدخلات.

س206/ هل المشفر التلقائي المتغير VAE أفضل من المشفر التلقائي AE؟

ج/ إذا كنت تريد تحكمًا دقيقًا في تمثيلاتك الكامنة وما تود أن تمثله ، فاختر VAE. ومع ذلك، إذا كانت AE كافية للعمل الذي تقوم به ، فما عليك سوى استخدام AE ، فهي بسيطة وغير معقدة بدرجة كافية.

س207/ ما هو العيب الأكثر شيوعاً في VAEs؟

ج/ أكبر عيب في المشفرات التلقائي المتغيرة VAEs هو أنها تميل إلى إنشاء مخرجات ضبابية **blurry** وغير واقعية **unrealistic**. يرتبط هذا بالطريقة التي يسترد بها VAE توزيعات البيانات وحساب دوال الخسارة (الخطأ).

س208/ ما هو المشفر التلقائي التلافيفي convolutional autoencoder؟

ج/ يعد المشفر التلقائي التلافيفي أحد أشكال الشبكات العصبية التلافيفية التي تُستخدم كأدوات للتعلم غير الخاضع للإشراف لمرشحات الالتفاف **convolution filters**. يتم تطبيقها بشكل عام في مهمة إعادة بناء الصورة لتقليل أخطاء إعادة البناء من خلال تعلم المرشحات المثلى.

س209/ هل شبكات GAN أفضل من VAE؟

ج/ أفضل شيء في VAE هو أنه يتعلم كل من النموذج التوليدي **generative** ونموذج الاستدلال **inference**. على الرغم من أن كلا من VAE و GAN هما نهجان مثيران للغاية لتعلم توزيع البيانات الأساسية باستخدام التعلم غير الخاضع للإشراف، إلا أن شبكات GAN تحقق نتائج أفضل مقارنةً بـ VAE.

س210/ هل يمكن استخدام المشفرات التلقائية لتقليل الضوضاء ؟

ج/ نعم، يمكن ذلك. في حالة المشفر التلقائي لإزالة الضوضاء **Denoising Autoencoder**، تكون البيانات تالفة جزئيًا بسبب الضوضاء المضافة إلى متجه الإدخال بطريقة عشوائية. بعد ذلك، يتم تدريب النموذج على التنبؤ بنقطة البيانات الأصلية غير التالفة كمخرجات لها.

س211/ ما هو المشفر التلقائي لازالة الضوضاء Denoising Autoencoder؟

ج/ يعد المشفر التلقائي لإزالة الضوضاء نوعًا محددًا من المشفر التلقائي، والذي يُصنف عمومًا على أنه نوع من الشبكات العصبية العميقة. يتم تدريب المشفر التلقائي لإزالة الضوضاء على استخدام طبقة مخفية لإعادة بناء نموذج معين بناءً على مدخلاته.

س212/ ما هو الفرق الرئيسي بين المشفر التلقائي Autoencoder والمشفر التلقائي لإزالة الضوضاء Denoising Autoencoder؟

ج/يقوم المشفر التلقائي لازالة الضوضاء بالتشفير، بالإضافة إلى تعلم ضغط البيانات (مثل المشفر التلقائي)، فهو يتعلم إزالة الضوضاء في الصور ، مما يسمح بأداء جيد حتى عندما تكون المدخلات صاخبة. لذا فإن المشفر التلقائي لازالة الضوضاء هو أكثر قوة من المشفر التلقائي , فهم يتعلمون المزيد من الميزات من البيانات أكثر من المشفر التلقائي القياسي.

س213/ ما هو المشفر التلقائي المتناثر sparse autoencoder ؟

ج/ يعد المشفر التلقائي المتناثر **sparse autoencoder** أحد أنواع الشبكات العصبية الاصطناعية للمشفر التلقائي التي تعمل على مبدأ التعلم الآلي غير الخاضع للإشراف. تعد المشفرات التلقائية نوعًا من الشبكات العميقة التي يمكن استخدامها لتقليل الأبعاد – ولإعادة بناء نموذج من خلال الانتشار العكسي.

س214/ هل يمكن استخدام المشفرات التلقائي للتجميع clustering؟

ج/في بعض جوانب ترميز **encoding** البيانات وتجميع **clustering** البيانات تشترك في بعض النظريات المتداخلة. نتيجة لذلك، يمكنك استخدام **Autoencoders** لتجميع (ترميز) البيانات. من الأمثلة البسيطة التي يمكنك تخيلها إذا كان لديك مجموعة من بيانات التدريب التي تشك في أنها تحتوي على فصلين أساسيين.

س215/ هل يمكن استخدام المشفرات التلقائي للتصنيف classification؟

ج/ نهج التشفير التلقائي للتصنيف مشابه لاكتشاف الشذوذ **anomaly detection**. في اكتشاف الشذوذ، نتعلم نمط العملية الطبيعية. يتم تصنيف أي شيء لا يتبع هذا النمط على أنه شذوذ **anomaly**.

س216/ هل RTX 3080 جيد للتعلم العميق؟

ج/ يعتبر **RTX 3080** أيضًا وحدة معالجة رسوميات ممتازة للتعلم العميق. ومع ذلك، فإن لها قيودًا واحدة وهي حجم **VRAM**. سيتطلب التدريب على **RTX 3080** أحجام دفعات صغيرة، لذلك قد لا يتمكن أولئك الذين لديهم طرز أكبر من تدريبهم.

س217/ أي GPU هو الأفضل للتعلم العميق؟

ج/ تعد **RTX 3090** من **NVIDIA** أفضل وحدة معالجة رسومات **GPU** للتعلم العميق والذكاء الاصطناعي. يتمتع بأداء استثنائي وميزات تجعله مثاليًا لتشغيل أحدث جيل من الشبكات العصبية. سواء كنت عالم بيانات أو باحثًا أو مطورًا، سيساعدك **RTX 3090** في الارتقاء بمشاريعك إلى المستوى التالي.

س218/ هل RTX أو GTX أفضل للتعلم العميق؟

ج/ **Nvidia** تكشف أن **RTX 2080 Ti** هو ضعف سرعة **GTX 1080 Ti**. ... نوى **Tensor** في كل وحدة معالجة رسومات **RTX** قادرة على أداء معالجة شبكة عصبية سريعة التعلم العميق للغاية وتستخدم هذه التقنيات لتحسين أداء اللعبة وجودة الصورة.

س219/ ما هي بطاقة الرسوميات CUDA؟

ج/ **CUDA** عبارة عن منصة حوسبة متوازية **parallel computing platform** ونموذج برمجة **programming model** طورته **Nvidia** للحوسبة العامة على وحدات معالجة الرسومات الخاصة بها (وحدات معالجة الرسومات). يُمكّن **CUDA** المطورين من تسريع التطبيقات كثيفة الحوسبة من خلال تسخير قوة وحدات معالجة الرسومات للجزء القابل للتوازي من الحساب.

س220/ ما هي نوى CUDA للتعلم العميق؟

ج/ للتعلم العميق، تُفضل نوى **CUDA** الخاصة بـ **Nvidia**، وبرامج تشغيل الرسومات مقارنة بوحدات المعالجة المركزية، لأن هذه النوى مصممة خصيصًا لمهام مثل المعالجة المتوازية، ورفع مستوى الصورة في الوقت الفعلي، وإجراء بيتافلوبات **petaflops** من العمليات الحسابية كل ثانية ، وعرض الفيديو عالي الدقة ، الترميز وفك الترميز.

س221/ ما هو البيتافلوب petaflops؟

ج/ بيتافلوب هو مقياس لسرعة معالجة الكمبيوتر. وحدة من سرعة الحوسبة تساوي ألف مليون مليون (10^{15}) عملية فاصلة عائمة في الثانية **floating-point operations per second**.

س222/ ما هو CUDA في CNN؟

ج/ مكتبة **NVIDIA CUDA® Deep Neural Network (cuDNN)** عبارة عن مكتبة مُسرعة من خلال وحدة معالجة الرسومات (**GPU**) من العناصر الأولية للشبكات العصبية العميقة. يوفر **cuDNN** تطبيقات دقيقة للغاية للإجراءات القياسية مثل الالتفاف الأمامي والخلفي **forward and backward convolution** ، والتجميع **pooling**، والتسوية **normalization** ، وطبقات التنشيط **activation layers**.

س223/ ما هو التعلم المعزز reinforcement learning؟

ج/التعلم المعزز هو نوع من أساليب التعلم الآلي حيث يتفاعل الوكيل الذكي (برنامج الكمبيوتر) مع البيئة ويتعلم التصرف ضمن ذلك. يتم تحديد أي مهمة تعلم معزز من خلال ثلاثة أشياء – الحالات **states** والإجراءات **actions** والمكافآت **rewards**. الحالات هي تمثيل للعالم الحالي أو بيئة المهمة. الإجراءات هي شيء يمكن لوكيل **RL** القيام به لتغيير هذه الحالات.

س224/ ما هي الحالة state والاجراء action في التعلم المعزز؟

ج/تتعلم خوارزمية التعلم المعزز، والتي قد يشار إليها أيضًا باسم الوكيل **agent**، من خلال التفاعل مع بيئتها **environment**. يتلقى الوكيل مكافآت **rewards** من خلال الأداء الصحيح والعقوبات **penalties** على الأداء غير الصحيح. يتعلم الوكيل دون تدخل من الإنسان من خلال زيادة مكافأته وتقليل العقوبة.

س225/ ما هي المكافأة reward والعقوبة penalty في التعلم المعزز؟

ج/تتعلم خوارزمية التعلم المعزز، والتي قد يشار إليها أيضًا باسم الوكيل **agent**، من خلال التفاعل مع بيئتها **environment**. يتلقى الوكيل مكافآت **rewards** من خلال الأداء الصحيح والعقوبات **penalties** على الأداء غير الصحيح. يتعلم الوكيل دون تدخل من الإنسان من خلال زيادة مكافأته وتقليل العقوبة.

س226/ ما هو المقصود بـ Q-Learning؟

ج/ **Q-Learning** هي خوارزمية تعلم معززة خالية من النموذج **model-free** لتعلم قيمة إجراء ما في حالة معينة. لا يتطلب نموذجًا للبيئة (وبالتالي "خالٍ من النماذج

model-free ") ، ويمكنه التعامل مع مشاكل التحولات العشوائية **stochastic transitions** والمكافآت **rewards** دون الحاجة إلى التكيف **adaptations**.

س227/ ما هي قيمة Q في التعلم المعزز؟

ج/ **Q-Learning** هو شكل أساسي من أشكال التعلم المعزز الذي يستخدم قيم **Q** (وتسمى أيضًا قيم الإجراء **action values**) لتحسين سلوك وكيل التعلم **learning agent** بشكل متكرر. قيم **Q** أو قيم الإجراءات **Action-Values**: يتم تعريف قيم **Q** للحالات **states** والإجراءات **actions**. هو تقدير لمدى جودة اتخاذ الإجراء في الحالة.

س228/ ما الفرق بين Q-Learning و Sarsa؟

ج/ تقنية **Q-Learning** هي تقنية خارج السياسة **Off Policy** وتستخدم النهج الجشع لتعلم قيمة **Q**. من ناحية أخرى ، فإن تقنية **SARSA** هي سياسة **On Policy** وتستخدم الإجراء الذي تقوم به السياسة الحالية لمعرفة قيمة **Q**.

س229/ ما هي أمثلة التعلم المعزز reinforcement learning؟

ج/ يمكننا القول أن "التعلم المعزز هو نوع من أساليب التعلم الآلي حيث يتفاعل الوكيل الذكي (برنامج الكمبيوتر) مع البيئة ويتعلم التصرف ضمن ذلك." كيف يتعلم الكلب الآلي حركة ذراعيه هو مثال على التعلم المعزز.

س230/ هل التعلم المعزز جزء من الذكاء الاصطناعي AI؟

ج/ إنه شكل من أشكال التعلم الآلي **ML** وبالتالي فرع من الذكاء الاصطناعي **AI**.

س231/ هل يحتاج التعلم المعزز إلى بيانات تدريبية trained data؟

ج/ يختلف التعلم المعزز عن الأساليب السابقة في أنه لا يحتاج إلى بيانات تدريبية ، ولكنه يعمل ويتعلم ببساطة من خلال نظام المكافآت **reward system** الموصوف.

س232/ ما هو الاختلاف الرئيسي بين التعلم العميق DL والتعلم المعزز RL؟

ج/ الفرق الرئيسي بينهما هو أن التعلم العميق يتعلم من مجموعة تدريب ثم يطبق هذا التعلم على مجموعة بيانات جديدة، بينما التعلم المعزز يتعلم ديناميكيًا عن طريق ضبط الإجراءات **actions** القائمة على ردود الفعل المستمرة **continuous feedback** لزيادة المكافأة **reward**.

س233/ ما هي عيوب التعلم المعزز؟

ج/ الكثير من التعلم المعزز يمكن أن يؤدي إلى عبء زائد من الحالات **overload of states**، مما قد يقلل من النتائج. لا يفضل استخدام التعلم المعزز لحل المشكلات البسيطة. يحتاج التعلم المعزز إلى الكثير من البيانات والكثير من الحسابات. إنها متعطشة للبيانات **data-hungry**.

س234/ هل التعلم المعزز مستخدم في الصناعة industry؟

ج/ في تعزيز الصناعة **industry reinforcement**، يتم استخدام الروبوتات القائمة على التعلم لأداء مهام مختلفة. بصرف النظر عن حقيقة أن هذه الروبوتات أكثر كفاءة من البشر، يمكنها أيضًا أداء مهام قد تكون خطرة على الناس.

س235/ ما هي التطبيقات واقعية للتعلم المعزز

ج/

- السيارات الذاتية القيادة
- أتمتة الصناعة
- التجارة والتمويل
- معالجة اللغة الطبيعية
- الرعاية صحية
- الهندسة
- التوصية الإخبارية
- الألعاب
- التسويق والدعاية
- الروبوتات

س236/ ما هو المقصود بالتعلم المعزز العميق deep reinforcement learning؟

ج/ يجمع التعلم المعزز العميق (**DRL**) بين التعلم المعزز والتعلم العميق. إنه أكثر قدرة على التعلم من أجهزة الاستشعار الخام أو الصور كمدخلات ، مما يتيح التعلم الشامل **end-to-end learning** ، والذي يفتح المزيد من التطبيقات في مجال الروبوتات وألعاب الفيديو ومعالجة اللغة الطبيعية ورؤية الكمبيوتر والرعاية الصحية والمزيد.

س237/ ما هو التعلم المعزز العميق وما هي تطبيقاته المختلفة؟

ج/ التعلم المعزز العميق هو فئة من التعلم الآلي والذكاء الاصطناعي حيث يمكن للآلات الذكية أن تتعلم من أفعالها على غرار الطريقة التي يتعلم بها البشر من التجربة. متأصل في هذا النوع من التعلم الآلي هو أن العامل او الوكيل يكافأ أو يُعاقب على أساس أفعاله.

س238/ ما هو الفرق بين التعلم المعزز والتعلم المعزز العميق؟

ج/ التعلم المعزز هو التعلم الديناميكي باستخدام طريقة التجربة والخطأ **trial and error** لتعظيم النتائج، بينما التعلم المعزز العميق يتعلم من المعرفة الحالية **existing knowledge** ويطبقها على مجموعة بيانات جديدة **new data set**.

س239/ ما هو التعلم العميق الهندسي Geometric Deep Learning؟

ج/ التعلم العميق الهندسي هو أحد المجالات المتخصصة في التعلم العميق الذي يهدف إلى تعميم نماذج الشبكة العصبية على المجالات غير الإقليدية مثل الرسوم البيانية **graphs** والتشعبات **manifolds**. مفهوم العلاقات والصلات والخصائص المشتركة هو مفهوم يحدث بشكل طبيعي في البشر والطبيعة.

التعلم العميق الهندسي هو مجال جديد للتعلم الآلي يمكنه التعلم من البيانات المعقدة مثل الرسوم البيانية والنقاط متعددة الأبعاد **multi-dimensional points**. يسعى إلى تطبيق الشبكات العصبية التلافيفية التقليدية على كائنات ثلاثية الأبعاد ورسوم بيانية ومشعبات **manifolds**.

س240/ ما هو التعلم العميق Q (Deep Q Learning)؟

ج/ يستخدم **Deep Q Learning** فكرة **Q-Learning** ويأخذها خطوة إلى الأمام. بدلاً من استخدام جدول **Q**، نستخدم شبكة عصبية تأخذ حالة وتقارب قيم **Q** لكل إجراء بناءً على تلك الحالة.

س241/ ما هو التعلم الانتقالي العميق Deep transfer learning (DTL)؟

ج/ تعد أساليب التعلم العميق للنقل (**DTL**) هي الحل لمعالجة قيود قلة البيانات، على سبيل المثال، أثبت ضبط نموذج مدرب مسبقًا على مجموعة بيانات ضخمة شبه ذات صلة أنه طريقة بسيطة وفعالة للعديد من المشكلات. تتعامل **DTLs** مع مخاوف البيانات المستهدفة المحدودة بالإضافة إلى تقليل تكاليف التدريب بشكل كبير.

س242/ ما هو التعلم الشامل end to end learning في التعلم العميق؟

ج/ التعلم الشامل في سياق الذكاء الاصطناعي والتعلم الآلي هو أسلوب يتعلم فيه النموذج جميع الخطوات بين مرحلة الإدخال الأولية والنتيجة النهائية للمخرجات. هذه عملية تعلم عميقة حيث يتم تدريب جميع الأجزاء المختلفة في وقت واحد بدلاً من التدريب المتسلسل.

س243/ ماذا تعني قاعدة دلتا Delta Rule في التعلم الآلي والعميق؟

ج/ تعتبر قاعدة دلتا في التعلم الآلي وبيئات الشبكة العصبية نوعًا محددًا من الانتشار الخلفي **backpropagation** الذي يساعد على تحسين شبكات **ML / AI** المتصلة ، مما يجعل الاتصالات بين المدخلات والمخرجات مع طبقات من الخلايا العصبية الاصطناعية. تُعرف قاعدة دلتا أيضًا بقاعدة تعليم دلتا **Delta learning rule**.

س244/ ما الخطوة التالية للتعلم الآلي؟

ج/ يمكن للحوسبة الكمومية **Quantum computing** أن تحدد مستقبل التعلم الآلي

س245/ ما الحوسبة الكمومية Quantum computing؟

ج/ تعتمد الحوسبة الكمومية على خصائص ميكانيكا الكم **quantum mechanics** لحساب المشكلات التي ستكون بعيدة المنال بالنسبة لأجهزة الكمبيوتر الكلاسيكية. يستخدم الكمبيوتر الكمومي الكيوبتات **qubits**. الكيوبتات تشبه البتات العادية في الكمبيوتر ، ولكن مع القدرة الإضافية على وضعها في التراكب ومشاركة التشابك مع بعضها البعض.

س246/ هل يمكن استخدام أجهزة الكمبيوتر الكمومية للتعلم العميق؟

ج/ يمكن دمج الحوسبة الكمومية مع التعلم العميق لتقليل الوقت اللازم لتدريب الشبكة العصبية. من خلال هذه الطريقة، يمكننا تقديم إطار عمل جديد للتعلم العميق وأداء التحسين الأساسي. يمكننا محاكاة خوارزميات التعلم العميق الكلاسيكية على كمبيوتر كمي حقيقي في العالم الحقيقي.

س247/ هل تعلم الآلة الكمي quantum machine learning حقيقي؟

ج/ نعم التعلم الآلي الكمي حقيقي ولكنه يظل مجالًا نظريًا بحتًا للدراسات.

س248/ ما المقصود بالتعلم الآلي الكمي quantum machine learning ؟

ج/ التعلم الآلي الكمي هو تكامل الخوارزميات الكمومية مع برامج التعلم الآلي. يشير الاستخدام الأكثر شيوعًا للمصطلح إلى خوارزميات التعلم الآلي لتحليل البيانات الكلاسيكية التي يتم تنفيذها على جهاز كمبيوتر كمي، أي التعلم الآلي المعزز بميكانيكا الكم. وبينما تُستخدم خوارزميات التعلم الآلي للتعامل مع

كميات هائلة من البيانات، فإن التعلم الآلي الكمي يستخدم الكيوبتات والعمليات الكمومية أو يستخدم أنظمة كمومية متخصصة لتحسين السرعة الحاسوبية وتخزين البيانات الذي تقوم به خوارزميات برنامج ما.

س249/ ما هو الجيل القادم من الذكاء الاصطناعي؟

سيكون الجيل التالي من الشبكات قادرًا على الإحساس والحساب والتعلم والعقل والتصرف وفقًا لنوايا العمل بشكل مستقل تقريبًا، وإدارة الانفجار المستمر للبيانات من عدد متزايد من الأجهزة الذكية المتصلة.

س250/ ما هي الأنواع الثلاثة للذكاء الاصطناعي؟

ج/هناك ثلاثة أنواع من الذكاء الاصطناعي (AI): الذكاء الاصطناعي الضيق **narrow AI** أو الضعيف **weak AI**، والذكاء الاصطناعي العام **general AI** أو القوي **strong AI** ، والذكاء الاصطناعي الخارق **super AI**. لقد حققنا حاليًا ذكاءً اصطناعيًا ضيقًا فقط.

س251/ ما الذي يرمز إليه الذكاء الاصطناعي العام GAI في الذكاء الاصطناعي؟

ج/ الذكاء الاصطناعي العام (GAI) هو تمثيل للقدرات المعرفية البشرية المعممة في البرامج بحيث يمكن لنظام **GAI** إيجاد حل عند مواجهة مهمة غير مألوفة. الغرض من نظام الذكاء الاصطناعي العام هو أداء أي مهمة يستطيع الإنسان القيام بها.

س252/ ما الفرق بين الذكاء الاصطناعي الضيق Narrow AI والذكاء الاصطناعي العام General AI؟

ج/ يشير الذكاء الاصطناعي الضيق إلى ما وصل إليه الذكاء الاصطناعي اليوم، بينما يشير الذكاء الاصطناعي العام إلى المكان الذي سيكون عليه في المستقبل. يُعرف أيضًا باسم الذكاء الاصطناعي العام (AGI) والذكاء الاصطناعي القوي **Strong AI**، وهو نوع من الذكاء الاصطناعي يمكنه التفكير والعمل تمامًا كما يفعل البشر.

س253/ ما هو الديب مايند DeepMind؟

ج/ **DeepMind** هو قسم من شركة **Alphabet**، مسؤول عن تطوير تكنولوجيا الذكاء الاصطناعي للأغراض العامة (AGI). تُعرف هذه التقنية أيضًا باسم **Google DeepMind**. يستخدم **DeepMind** بيانات البكسل الأولية كمدخلات ويتعلم من التجربة.

س254/ ما هي البرامج التي يستخدمها DeepMind؟

ج/ على الرغم من أن **Google** لا تشارك جميع تفاصيل أسرار التعلم الآلي الخاصة بها، إلا أن الكثير من التعليمات البرمجية الخاصة بها مفتوحة المصدر، وقد شاركت أيضًا رمز برنامجها المسمى **TensorFlow**، وهو محرك التعلم العميق. تنشر **DeepMind** أيضًا العديد من الأوراق الأكاديمية حول عملها عبر الإنترنت.

س255/ لماذا يحظى التعلم العميق بشعبية في السنوات الأخيرة؟

ج/ في الآونة الأخيرة، اكتسب التعلم العميق شعبية كبيرة بسبب تفوقه من حيث الدقة عند تدريبه على كمية هائلة من البيانات. تتجه صناعة البرمجيات الآن بعد أيام نحو ذكاء الآلة. أصبح التعلم الآلي ضروريًا في كل قطاع كوسيلة لجعل الآلات ذكية.

س256/ ما هو مستقبل التعلم العميق ؟

ج/ بعد بضع سنوات من الآن، من المحتمل جدًا أن تصبح أدوات تطوير التعلم العميق والمكتبات واللغات مكونات قياسية لكل مجموعة أدوات لتطوير البرامج. ستمهد مجموعات الأدوات هذه ذات الإمكانات الحديثة الطريق لسهولة تصميم النماذج الجديدة وتكوينها وتدريبها

256

Question & Answers

In

ML, DL and DS

By:
Dr. Alaa Taima