الدورة التدريبية: تقنيات زيادة البيانات لنماذج الذكاء الاصطناعي القوية

يوليو ٢٠٢٦ ٣١ - ٢٧

ميونخ

(للشخص الواحد) € ٥١٠٠

Ref: #AI4596_247908

مقدمة الدورة التدريبية / لمحة عامة:

القوية، وهي مصممة لعلماء التدريبية المتخصصة حول تقنيات زيادة البيانات يقدم BIG BEN Training Center هذه الدورة والمطورين الذين يعملون في مجال الذكاء الاصطناعي البيانات، ومهندسي التعلم الآلي، والباحثين، لنماذج الذكاء الاصطناعي (AI)، أو الحاجة ويواجهون تحديات نقص البيانات (Data Scarcity)، ،(Deep Learning) والتعلم العميق ((Data Imbalance فعال، مما يساعد تقنية حيوية لتوسيع (Data Augmentation) إلى تحسين أداء النماذج. تُعد زيادة البيانات أو توازن البيانات الانحياز (Bias))، وزيادة قوة نماذج على تحسين تعميم النماذج (Model Generalization)، مجموعات البيانات التدريبية بشكل (Natural Language)، اللغات ، ومعالجة (Image Classification) الذكاء الاصطناعي، خاصةً في مهام مثل تصنيف الصور وتقليل الدورة إلى تزويد المشاركين بالمعرفة والمهارات والرؤية الحاسوبية (.(Computer Vision تهدف الطبيعية (Processing - NLP) Transformations) البيانات**، بدءاً من التحويلات البسيطة (Simple) اللازمة لتطبيق مجموعة واسعة من تقنيات زيادة) لتوليد بيانات اصطناعية المدعومة بالذكاء الاصطناعي مثل الشبكات التوليدية وصولاً إلى الأساليب المتقدمة في الصور؛ البيانات في الذكاء الاصطناعي، وأنواع زيادة ستغطي الدورة مفاهيم أهمية .(Synthetic Data) التنافسية (GANs) لكل من البيانات المرئية والمرادفات، وإعادة الصياغة في النصوص)، وتقنيات البيانات** (مثل التدوير، والقلب، والقص البيانات المتخصصة

عمل شائعة مثل تقنيات زيادة البيانات المناسبة لمشاكل تعلم آلي والنصية والسمعية. سيتعلم المشاركون كيفية اختيار زيادة على أداء النموذج. تهدف الدورة إلى تمكين ، وتقييم تأثير زيادة PyTorch و TensorFlow معينة، وتطبيقها باستخدام أطر الدورة ذكاء اصطناعي أكثر قوة ومرونة، وتحقيق نتائج أفضل المختصين من التغلب على قيود البيانات، وبناء نماذج البيانات أكد على Yann LeCun من أعمال البروفيسور يان ليكون (Yann LeCun) ، في مشاريع الذكاء الاصطناعي. نستلهم في هذه فعالة. أهمية البيانات الكافية والمتنوعة لتدريب نماذج ، أحد رواد التعلم العميق، الذي

لـ: الفئات المستهدفة / هذه الدورة التدريبية مناسبة

- علماء البيانات.
- مهندسي التعلم الآلي.
- مهندسي التعلم العميق.
- الباحثين في مجال الذكاء الاصطناعي.
- مطورين الذكاء الاصطناعي.
- الاصطناعي. طلاب الدراسات العليا في علوم الحاسوب والذكاء
- المهتمين بتحسين أداء نماذج التعلم العميق.
- المتخصصين في الرؤية الحاسوبية.
- المتخصصين في معالجة اللغات الطبيعية.
- المختصين في تحليل البيانات الكبيرة.

القطاعات والصناعات المستهدفة:

- تكنولوجيا المعلومات والبرمجيات.
- البحث والتطوير (R&D).
- الرعاية الصحية (توليد بيانات طبية).
- السيارات ذاتية القيادة (بيانات القيادة).
- الأمن والمراقبة.
- الزراعة الذكية.
- الخدمات المالية (بيانات المعاملات).
- التجارة الإلكترونية (بيانات المنتجات والصور).
- الألعاب (توليد الأصول).
- التعليم والبحث الأكاديمي.

الأقسام المؤسسية المستهدفة:

- قسم علوم البيانات.
- قسم التعلم الآلي والذكاء الاصطناعي.
- قسم البحث والتطوير.
- قسم هندسة البرمجيات (فرق الذكاء الاصطناعي).
- قسم الابتكار.
- قسم تحليل البيانات.
- فريق تطوير المنتجات القائمة على الذكاء الاصطناعي.
- الفرق المتخصصة في الرؤية الحاسوبية.
- الفرق المتخصصة في معالجة اللغات الطبيعية.
- وحدات المحاكاة.

أهداف الدورة التدريبية:

أتقن المهارات التالية: بنهاية هذه الدورة التدريبية، سيكون المتدرب قد

- الذكاء الاصطناعي. فهم أهمية زيادة البيانات في تحسين أداء نماذج
- للصور. التعرف على أنواع مختلفة من تقنيات زيادة البيانات
- بيانات الصور. تطبيق تحويلات هندسية ولونية لـزيادة مجموعات
- الكلمات والمرادفات). استخدام تقنيات زيادة البيانات للنصوص (مثل تبديل
- لـتوليد بيانات اصطناعية. تطبيق الشبكات التوليدية التنافسية ((GANs
- التعامل مع مشكلة نقص البيانات وعدم توازن الفئات.
- محددة. اختيار أفضل استراتيجية لـزيادة البيانات لمشكلة
- شائعة. Python تنفيذ تقنيات زيادة البيانات باستخدام مكتبات
- تقييم تأثير زيادة البيانات على أداء النموذج.
- بناء نماذج تعلم عميق أكثر قوة ومرونة.

منهجية الدورة التدريبية:

المكثف، بهدف منهجية تدريبية تجمع بين الفهم النظري لـتقنيات يعتمد BIG BEN Training Center في هذه الدورة على البيانات المحدودة. تشمل المنهجية تمكين المشاركين من بناء نماذج ذكاء اصطناعي قوية زيادة البيانات والتطبيق العملي البيانات، وتحديات نقص البيانات في التعلم العميق، محاضرات تفاعلية تستعرض المفاهيم الأساسية لزيادة حتى مع مجموعات تقنيات زيادة البيانات تتبع هذه المحاضرات ورش عمل تطبيقية مكثفة حيث والأساليب التقليدية والمتقدمة لزيادة البيانات. ، ومكتبات مثلPyTorch و TensorFlow على بيانات الصور والنصوص باستخدام أطر عمل مثل سيقوم المشاركون بـتطبيق نماذج كيفية استخدام زيادة البيانات لحل مشكلات مثل تصنيف و .NLTKسيتم التركيز على دراسات حالة واقعية تبرز OpenCV استراتيجيات زيادة بيانات مخصصة معالجة اللغات الطبيعية. تتضمن الدورة جلسات عمل الصور الطبية ببيانات قليلة، أو تحسين جودة المشاركون تغذية راجعة مفصلة ومنتظمة من المدربين لمشاريع الذكاء الاصطناعي المختلفة. يتلقى جماعي لـتطوير البيانات وأداء النماذج. الخبراء لضمان تطوير مهاراتهم في تحسين

خريطة المحتوى التدريبي (محاور الدورة التدريبية):

الوحدة الأولى: مقدمة إلى زيادة البيانات وأهميتها.

- مشكلة نقص البيانات في التعلم العميق.
- لماذا نحتاج إلى زيادة البيانات؟
- الناقص. مفاهيم أساسية: التحيز، التعميم الزائد، التعميم
- أنواع البيانات: صور، نصوص، صوت.
- الهدف من زيادة البيانات: تحسين الأداء والمرونة.
- زيادة البيانات مقابل جمع المزيد من البيانات.
- أهمية جودة البيانات في بناء نماذج قوية.

الهندسية واللونية). الوحدة الثانية: زيادة البيانات للصور (التحويلات

- الترجمة. التحويلات الهندسية: التدوير، الانعكاس، القص،
- الضوضاء. التحويلات اللونية: تغيير السطوع، التباين، التشبع،
- تطبيق مجموعات من التحويلات لتوليد بيانات متنوعة.
- تقنيات المزيج (Mix-up) و Cut Mix.
- (Augmentations, Imgaug) استخدام مكتبات Python لزيادة البيانات المرئية
- دراسة حالة: زيادة البيانات لتصنيف الصور الطبية.
- أفضل الممارسات في زيادة بيانات الصور.

(باستخدام. (GANs الوحدة الثالثة: زيادة البيانات المتقدمة للصور

- (GANs) مقدمة سريعة إلى الشبكات التوليدية التنافسية
- اصطناعية. كيف يمكن للشبكات التوليدية التنافسية توليد بيانات
- تطبيق GANs لإنشاء صور واقعية لزيادة البيانات.
- تحديات تدريب GANs لزيادة البيانات.
- البيانات (Cycle GAN). أمثلة على معماريات GANs المستخدمة لزيادة
- (Conditional Data Generation) تقنيات توليد البيانات المشروطة
- التدريب. دمج البيانات الحقيقية والاصطناعية لتحسين

الجدولية.ا الوحدة الرابعة: زيادة البيانات للنصوص والبيانات

- تقنيات زيادة البيانات للنصوص:ا
- تبديل الكلمات (Word Swapping)، إضافة المرادفات.ا
- إعادة صياغة الجمل (Sentence Paraphrasing).ا
- زيادة البيانات السياقية (Contextual Augmentation).ا
- لزيادة البيانات النصية. (استخدام نماذج اللغة الكبيرة (Large Language Models
- زيادة البيانات للبيانات الجدولية:ا
- التوليد الاصطناعي للبيانات الجدولية.ا
- التعامل مع البيانات العددية والفئوية.ا
- الطبيعية.ا أمثلة على زيادة البيانات في تطبيقات معالجة اللغات

وتقييمها.ا الوحدة الخامسة: استراتيجيات زيادة البيانات

- اختيار استراتيجية زيادة البيانات المناسبة.ا
- تأثير حجم مجموعة البيانات على اختيار التقنيات.ا
- البيانات.ا التعامل مع عدم توازن الفئات باستخدام زيادة
- تقييم تأثير زيادة البيانات على أداء النموذج.ا
- مقاييس الأداء (الدقة، الاستدعاء، F1-Score).ا
- الزيادة التلقائية للبيانات (Auto Augment).ا
- الاصطناعي.ا الآفاق المستقبلية في زيادة البيانات والتعلم

الأسئلة المتكررة:ا

التسجيل في الدورة؟ ما هي المؤهلات أو المتطلبات اللازمة للمشاركين قبل

لا توجد شروط مسبقة.ا

الإجمالي لساعات الدورة التدريبية؟ كم تستغرق مدة الجلسة اليومية، وما هو العدد

المدة إلى ٢٥٢٠- بمعدل يومي يتراوح بين ٤ إلى ٥ ساعات، تشمل فترات تمتد هذه الدورة التدريبية على مدار خمسة أيام، ساعة تدريبية، راحة وأنشطة تفاعلية، ليصل إجمالي

سؤال للتأمل:

على هذه البيانات في توليد مجموعات بيانات اصطناعية واقعية بشكل متزايد، مع القدرة المتزايدة لتقنيات زيادة البيانات على هذا الاعتماد من الحاجة، إلى جمع بيانات تدريب نماذج الذكاء الاصطناعي الحساسة، وهل يمكن أن إلى أي، مدى يمكننا الاعتماد حقيقية مكلفة وعالية الجودة؟ يقلل

ما الذي يميز هذه الدورة عن غيرها من الدورات؟

من بناء نماذج ذكاء في مجال زيادة البيانات (Data Augmentation)، مع تتميز هذه الدورة بتقديمها نهجاً عملياً ومتقدماً دمج التقنيات التقليدية لزيادة البيانات مع اصطناعي قوية حتى في ظل قيود البيانات. ما يميزنا التركيز على تمكين المشاركين مجموعة واسعة من مثل استخدام الشبكات التوليدية التنافسية (GANs) الأساليب الحديثة القائمة على الذكاء الاصطناعي، هو لكيفية التعامل مع تحديات البيانات أنواع البيانات (صور، نصوص، جداول)، مما يوفر فهماً لتوليد بيانات اصطناعية**. نغطي مشكلة نقص البيانات، وتحقيق بالمهارات اللازمة لتحسين أداء نماذج التعلم المتنوعة. الدورة تركز على تزويد المشاركين شاملاً الذكاء الاصطناعي يسعى للارتقاء بقدراته في تعميم أفضل للنماذج، مما يجعلها ضرورية لأي متخصص العميق، والتغلب على معالجة البيانات وبناء النماذج. في