

Nama : Arfara Yema Samgusdian

NIM : 1103202004

Kelas : TK-44-04

Zero to Mastery Learn PyTorch for Deep Learning

08. PyTorch Paper Replicating

Apa itu Replikasi Paper dalam Machine Learning?

Replikasi paper dalam machine learning adalah usaha untuk mereplikasi kemajuan-kemajuan yang dihasilkan oleh penelitian machine learning dan diuraikan dalam sebuah paper. Tujuannya adalah agar kita dapat menggunakan teknik-teknik tersebut untuk memecahkan masalah kita sendiri.

Apa itu Paper Penelitian Machine Learning?

Paper penelitian machine learning adalah sebuah karya ilmiah yang merinci temuan dari suatu kelompok penelitian di bidang tertentu. Struktur umum dari paper penelitian machine learning melibatkan beberapa bagian:

Abstrak: Ringkasan dari temuan/kontribusi utama paper.

Pendahuluan: Deskripsi masalah utama yang diangkat dalam paper dan rincian metode-metode sebelumnya yang digunakan untuk mencoba memecahkan masalah tersebut.

Metode: Bagaimana para peneliti melakukan penelitian mereka? Contohnya, model(s), sumber data, dan setup pelatihan apa yang digunakan?

Hasil: Apa hasil dari paper tersebut? Jika model atau setup pelatihan baru digunakan, bagaimana hasilnya dibandingkan dengan penelitian sebelumnya? (di sinilah pelacakan eksperimen berguna).

Kesimpulan: Apa keterbatasan dari metode yang diusulkan? Apa langkah-langkah berikutnya untuk komunitas penelitian?

Referensi: Sumber daya atau paper lain yang digunakan oleh peneliti untuk membangun karyanya sendiri.

Lampiran: Apakah ada sumber daya/temuan tambahan yang dapat dilihat yang tidak termasuk dalam bagian-bagian di atas?

Mengapa Melakukan Replikasi Paper dalam Machine Learning?

Paper penelitian machine learning seringkali merupakan presentasi dari bulan-bulan atau bahkan tahun-tahun kerja dan eksperimen yang dilakukan oleh beberapa tim machine learning terbaik di dunia, yang dirangkum dalam beberapa halaman teks.

Melakukan eksperimen yang mengarah pada hasil yang lebih baik dalam bidang yang terkait dengan masalah yang Anda kerjakan, tentu akan menjadi hal yang baik untuk dicoba.

Di Mana Bisa Menemukan Contoh Kode untuk Paper Penelitian Machine Learning?

Ada beberapa tempat untuk menemukan dan membaca paper penelitian machine learning beserta contoh kode:

arXiv: Sumber daya gratis dan terbuka untuk membaca artikel teknis di berbagai bidang, termasuk machine learning.

AK Twitter: Akun Twitter AK mempublikasikan sorotan penelitian machine learning, seringkali dengan demonstrasi langsung hampir setiap hari.

Papers with Code: Kumpulan paper machine learning yang tren, aktif, dan terbaik, banyak di antaranya menyertakan sumber daya kode.

GitHub vit-pytorch milik lucidrains: Bukan tempat untuk menemukan paper penelitian, tetapi merupakan contoh replikasi paper dalam skala lebih besar dan dengan fokus tertentu.

Repository vit-pytorch adalah kumpulan arsitektur model Vision Transformer dari berbagai paper penelitian yang direplikasi dengan kode PyTorch.

0. Persiapan

Unduh modul Python yang diperlukan dan pastikan kita memiliki semua skrip yang dibutuhkan dari sebelumnya.

Dapatkan direktori going_modular dari repositori pytorch-deep-learning.

Unduh paket torchinfo dan pastikan torchvision v0.13.

1. Dapatkan Data

Unduh dataset gambar pizza, steak, dan sushi menggunakan fungsi `download_data()` dari `helper_functions.py`.

2. Buat Datasets dan DataLoaders

Gunakan fungsi `create_dataloaders()` dari `data_setup.py` untuk membuat DataLoaders.

Persiapkan transformasi untuk gambar, dengan resolusi training sebesar 224x224, mengacu pada ViT paper.

3. Replikasi Paper ViT: Gambaran Umum

Replikasi paper ViT untuk FoodVision Mini, mengonversi input gambar menjadi label kelas yang diprediksi.

Ide dasarnya adalah menggunakan lapisan, blok, dan fungsi untuk menyusun arsitektur keseluruhan.

4. Persamaan 1: Pisahkan Data menjadi Patches dan Buat Embeddings

Fokus pada membuat embedding untuk kelas, posisi, dan patch sesuai dengan Equation 1 dalam ViT paper.

Mulai dengan membuat patch embedding, mengubah gambar input menjadi urutan patch dan membuat embedding untuk setiap patch.

Gunakan `torch.nn.Conv2d()` untuk membagi gambar menjadi patch dan membuat embedding untuk setiap patch.

Input: Gambar 2D dengan ukuran $H \times W \times C$.

Output: Urutan 1D dari patch 2D yang diratakan dengan ukuran $N \times (P^2 \cdot C)$.

Langkah selanjutnya termasuk langkah-langkah berikutnya dalam replikasi paper ViT.