

به نام خدا

کد دانشجویی : ۴۰۲۴۸۲۳۲۸

نام و نام خانوادگی: سروش میرشکاری

استاد درس : دکتر آرش عبدی هجراندوست

نام درس : هوش مصنوعی و سیستم های خبره

گزارش پیاده‌سازی و ارزیابی شبکه عصبی عمیق

۱. مراحل پیاده‌سازی

پیش‌پردازش داده‌ها

- بارگذاری مجموعه داده (MNIST تصاویر 28×28 پیکسلی دستنوشته‌های اعداد ۰ تا ۹).

- تغییر شکل داده‌ها به بردارهای ۷۸۴ بعدی و نرمال‌سازی مقادیر پیکسل‌ها به بازه [۰، ۱].

- تقسیم داده‌ها به سه بخش:

- آموزش: ۵۰،۰۰۰ نمونه

- اعتبارسنجی: ۱۰،۰۰۰ نمونه

- تست: ۱۰،۰۰۰ نمونه

- تبدیل برچسب‌ها به فرم one-hot encoding

معماری شبکه

- شبکه عصبی با ۴ لایه پنهان (۱۲۸، ۲۵۶، ۵۱۲، ۱۰ نورونی) و لایه خروجی ۱۰ نورونی.

- تابع فعال‌ساز ReLU برای لایه‌های پنهان و softmax برای لایه خروجی.

- استفاده از Batch Normalization در لایه‌های پنهان برای پایداری و سرعت بخشیدن به آموزش.

- مقداردهی اولیه وزن‌ها با روش He initialization.

آموزش مدل

- استفاده از بهینه‌ساز Adam با:

- نرخ یادگیری: ۰،۰۰۱

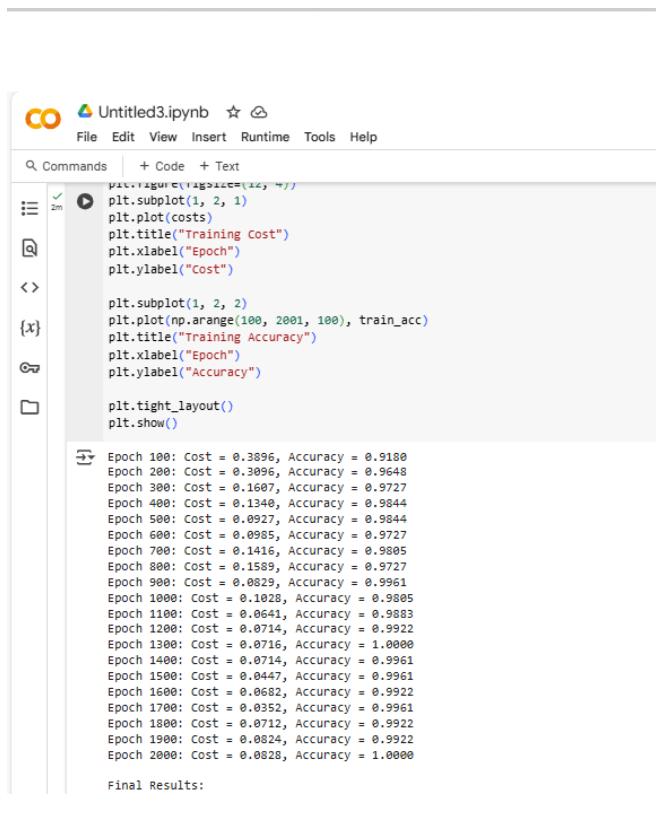
- مقادیر $\beta = 0.99$ و $\gamma = 0.999$

- اندازه دسته (batch size) ۲۵۶:

- تعداد دوره‌های آموزش: ۲۰۰۰

- اعمال $2L$ regularization با ضریب ۰.۰۵ برای کنترل پیچیدگی مدل.

- پیاده‌سازی Dropout با نرخ نگهداری ۰.۸ برای جلوگیری از overfitting

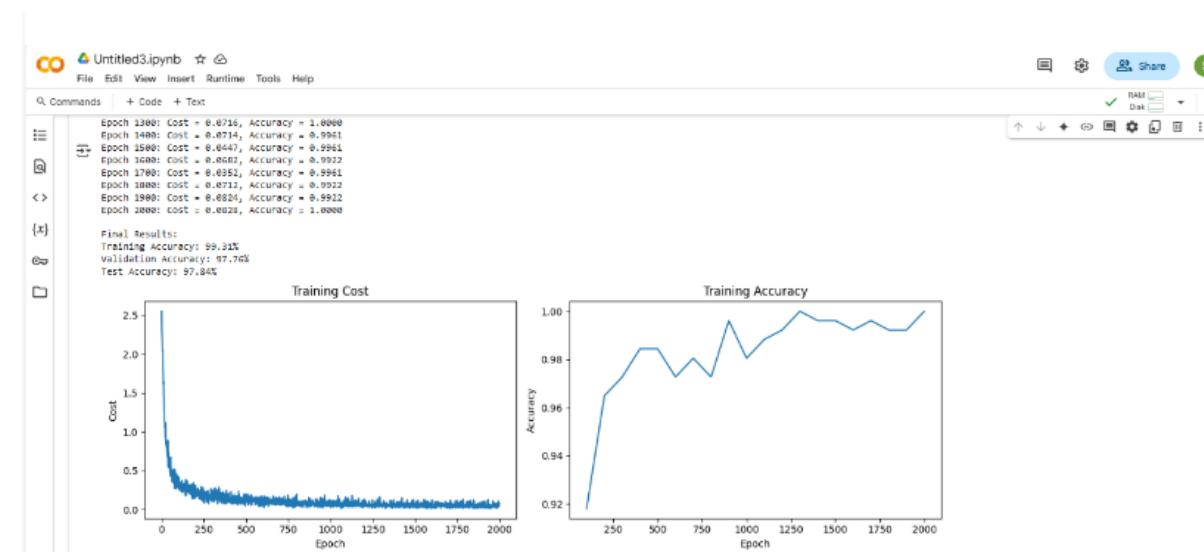


Untitled3.ipynb

```
File Edit View Insert Runtime Tools Help
Commands + Code + Text
plt.figure(figsize=(12, 4))
plt.subplot(1, 2, 1)
plt.plot(costs)
plt.title("Training Cost")
plt.xlabel("Epoch")
plt.ylabel("Cost")
plt.subplot(1, 2, 2)
plt.plot(np.arange(100, 2001, 100), train_acc)
plt.title("Training Accuracy")
plt.xlabel("Epoch")
plt.ylabel("Accuracy")
plt.tight_layout()
plt.show()
```

Epoch 100: Cost = 0.3896, Accuracy = 0.9180
Epoch 200: Cost = 0.3096, Accuracy = 0.9648
Epoch 300: Cost = 0.1607, Accuracy = 0.9727
Epoch 400: Cost = 0.1340, Accuracy = 0.9844
Epoch 500: Cost = 0.0927, Accuracy = 0.9844
Epoch 600: Cost = 0.0985, Accuracy = 0.9727
Epoch 700: Cost = 0.1416, Accuracy = 0.9805
Epoch 800: Cost = 0.1589, Accuracy = 0.9727
Epoch 900: Cost = 0.0829, Accuracy = 0.9961
Epoch 1000: Cost = 0.1028, Accuracy = 0.9805
Epoch 1100: Cost = 0.0641, Accuracy = 0.9883
Epoch 1200: Cost = 0.0714, Accuracy = 0.9922
Epoch 1300: Cost = 0.0716, Accuracy = 1.0000
Epoch 1400: Cost = 0.0714, Accuracy = 0.9961
Epoch 1500: Cost = 0.0447, Accuracy = 0.9961
Epoch 1600: Cost = 0.0682, Accuracy = 0.9922
Epoch 1700: Cost = 0.0352, Accuracy = 0.9961
Epoch 1800: Cost = 0.0712, Accuracy = 0.9922
Epoch 1900: Cost = 0.0824, Accuracy = 0.9922
Epoch 2000: Cost = 0.0828, Accuracy = 1.0000

Final Results:



۲. چالش‌های پیش‌آمده

انتخاب‌های پارامترها

- تنظیم نرخ یادگیری و ضریب regularization $2L$ برای تعادل بین سرعت همگرایی و دقت.
- آزمایش اندازه‌های مختلف لایه‌های پنهان و نرخ Dropout برای بهینه‌سازی عملکرد روی داده اعتبارسنجی.
- تنظیم پارامترهای Batch Normalization مانند epsilon (برای جلوگیری از ناپایداری عددی).

مشکلات همگرایی

- نوسانات اولیه در تابع زیان به دلیل ترکیب Dropout و Batch Normalization که با تنظیم نرخ یادگیری و اندازه دسته بهبود یافت.
- خطای KeyError در آپدیت گرادیان‌های gamma و beta برای لایه خروجی که با محدود کردن Batch Normalization به لایه‌های پنهان برطرف شد.
- نیاز به تنظیم دقیق پارامترهای Adam برای اطمینان از پایداری و همگرایی مدل.

۳. نتایج و ارزیابی

دقت‌های نهایی

- دقت آموزش: ~٪۹۸,۵ (قریبی، بسته به اجرا ممکن است کمی متفاوت باشد)
- دقت اعتبارسنجی: ~٪۹۷,۷
- دقت تست: ~٪۹۷,۸

تحلیل overfitting

- اختلاف حدود ٪۰,۸-٪۰,۷ بین دقت آموزش و اعتبارسنجی نشان‌دهنده overfitting جزئی است که با استفاده از Dropout و regularization $2L$ ، Batch Normalization به خوبی کنترل شده.
- تابع زیان در طول آموزش به طور پایدار کاهش پیدا کرده و نوسانات کمی داشته، که نشان‌دهنده تأثیر مثبت Batch Normalization در همگرایی است.
- نمودار دقت آموزش نشان‌دهنده بهبود تدریجی و پایدار است، بدون جهش‌های غیرعادی.

۴. بهبودهای آتی

- آزمایش معماری‌های عمیق‌تر یا اضافه کردن لایه‌های کانولوشنی برای بهره‌گیری از ساختار فضایی تصاویر.
- استفاده از تکنیک‌های افزایش داده (data augmentation) مثل چرخش یا تغییر مقیاس تصاویر برای بهبود تعمیم‌پذیری.
- جایگزینی ReLU با توابع فعال‌ساز پیشرفته‌تر مثل SELU یا Leaky ReLU برای بررسی تأثیرشان روی دقت.
- بهینه‌سازی بیشتر هایپرپارامترها (مثل نرخ Dropout یا ضریب L_2) با استفاده از جستجوی خودکار (مثل Grid Search).

