

به نام خدا

نام و نام خانوادگی: سروش میرشکاری

کد دانشجویی: ۴۰۲۴۸۲۳۲۸

نام درس: هوش مصنوعی و سیستم های خبره

استاد درس: دکتر آرش عبدی هجراندوست

گزارش پیاده سازی و ارزیابی شبکه عصبی عمیق

۱. مراحل پیاده سازی

پیش پردازش داده ها

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

- تغییر شکل داده ها به بردارهای ۷۸۴ بعدی و نرمال سازی مقادیر پیکسل ها به بازه $[0, 1]$.

- تقسیم داده ها به سه بخش:

- آموزش: ۵۰,۰۰۰ نمونه

- اعتبار سنجی: ۱۰,۰۰۰ نمونه

- تست: ۱۰,۰۰۰ نمونه

- تبدیل برچسب ها به فرم one-hot encoding

معماری شبکه

- شبکه عصبی با ۴ لایه پنهان (۵۱۲، ۲۵۶، ۱۲۸، و ۶۴ نورونی) و لایه خروجی ۱۰ نورونی.

- تابع فعال ساز ReLU برای لایه های پنهان و softmax برای لایه خروجی.

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

- مقداردهی اولیه وزن ها با روش He initialization.

آموزش مدل

- استفاده از بهینه ساز Adam با:

- نرخ یادگیری: ۰,۰۰۱

– مقادیر $\beta_1=0.9$ و $\beta_2=0.999$

– اندازه دسته (batch size): ۲۵۶

– تعداد دوره‌های آموزش: ۲۰۰۰

– اعمال L2 regularization با ضریب ۰.۰۵ برای کنترل پیچیدگی مدل.

– پیاده‌سازی Dropout با نرخ نگهداری ۰.۸ برای جلوگیری از overfitting

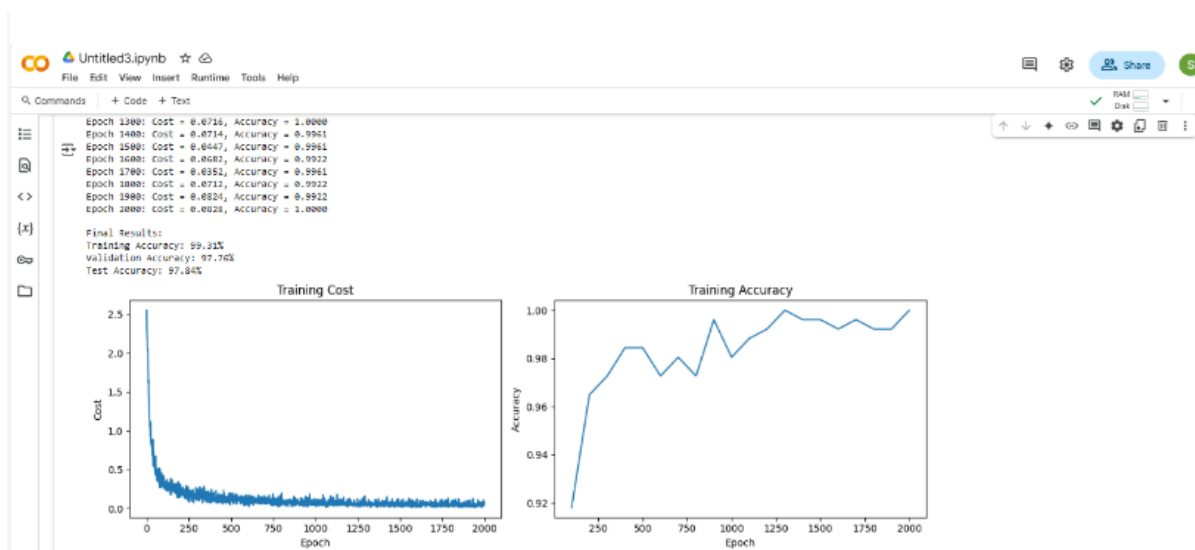
```
plt.figure(figsize=(12, 7))
plt.subplot(1, 2, 1)
plt.plot(costs)
plt.title("Training Cost")
plt.xlabel("Epoch")
plt.ylabel("Cost")

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 پیش آمده

انتخاب هایپر پارامترها

- تنظیم نرخ یادگیری و ضریب γ regularization 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 گرادینت‌های 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: $\sim 98.5\%$ (تقریبی، بسته به اجرا ممکن است کمی متفاوت باشد)
- دقت اعتبارسنجی: $\sim 97.7\%$.
- دقت تست: $\sim 97.8\%$.

تحلیل overfitting

- اختلاف حدود $0.7\% - 0.8\%$ بین دقت آموزش و اعتبارسنجی نشان‌دهنده overfitting جزئی است که با استفاده از Dropout، γ regularization و 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 Leaky ReLU یا SELU برای بررسی تأثیرشان روی دقت.
- بهینه‌سازی بیشتر هایپرپارامترها (مثل نرخ Dropout یا ضریب γ) با استفاده از جستجوی خودکار (مثل Grid Search).

