



به نام خدا

درس یادگیری ماشین

تمرین چهارم

حدیثه مصباح

۸۱۰۱۰۲۲۵۳



پاسخ ۵.

**بخش ۱:** این کد یک شبکه عصبی چندلایه (Multilayer Perceptron) را برای داده‌های MNIST ایجاد و آموزش می‌دهد. در اینجا توضیحاتی در مورد پارامترهای شبکه و تعداد لایه‌های مختلف ارائه شده است:

۱. لایه‌های شبکه:

- Flatten: این لایه تصاویر ورودی  $28 \times 28$  پیکسلی را به یک بردار یک‌بعدی ۷۸۴ عنصری تبدیل می‌کند.  
( $784 = 28 \times 28$ )

- Dense (۱۲۸ نورون, فعال‌سازی 'relu'): این لایه چگال (Dense) یک لایه متصل کاملاً به همه نورون‌های لایه قبلی است. این لایه ۱۲۸ نورون دارد و از تابع فعال‌سازی 'relu' (Rectified Linear Unit) استفاده می‌کند.

- Dense (۱۰ نورون, فعال‌سازی 'softmax'): این لایه نیز مانند لایه قبل چگال است اما تنها ۱۰ نورون دارد که هر کدام متناظر با یکی از ۱۰ کلاس دسته‌بندی (اعداد ۰ تا ۹) هستند. تابع فعال‌سازی 'softmax' احتمال تعلق هر تصویر ورودی به هر یک از کلاس‌ها را محاسبه می‌کند.

۲. پارامترها:

- تابع هزینه ('categorical\_crossentropy'): این تابع برای مسائل دسته‌بندی چند کلاسی مناسب است.

- بهینه‌ساز Adam ('adam'): یک الگوریتم بهینه‌سازی پیشرفته است که از مفاهیم مومنوم و تطبیقی نرخ یادگیری استفاده می‌کند.

- معیار ('accuracy'): این معیار دقت طبقه‌بندی را اندازه‌گیری می‌کند.

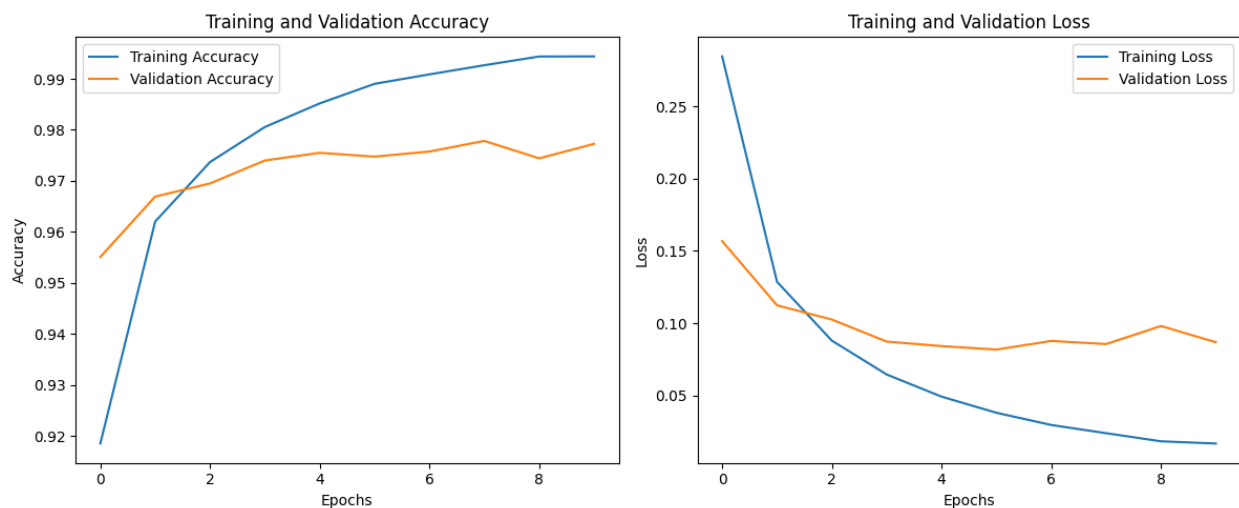
۳. آموزش:

- شبکه با استفاده از داده‌های آموزشی برای ۱۰ دوره (epoch) آموزش داده می‌شود.

- همچنین از داده‌های اعتبارسنجی (validation data) برای ارزیابی عملکرد مدل در هر دوره استفاده می‌شود.



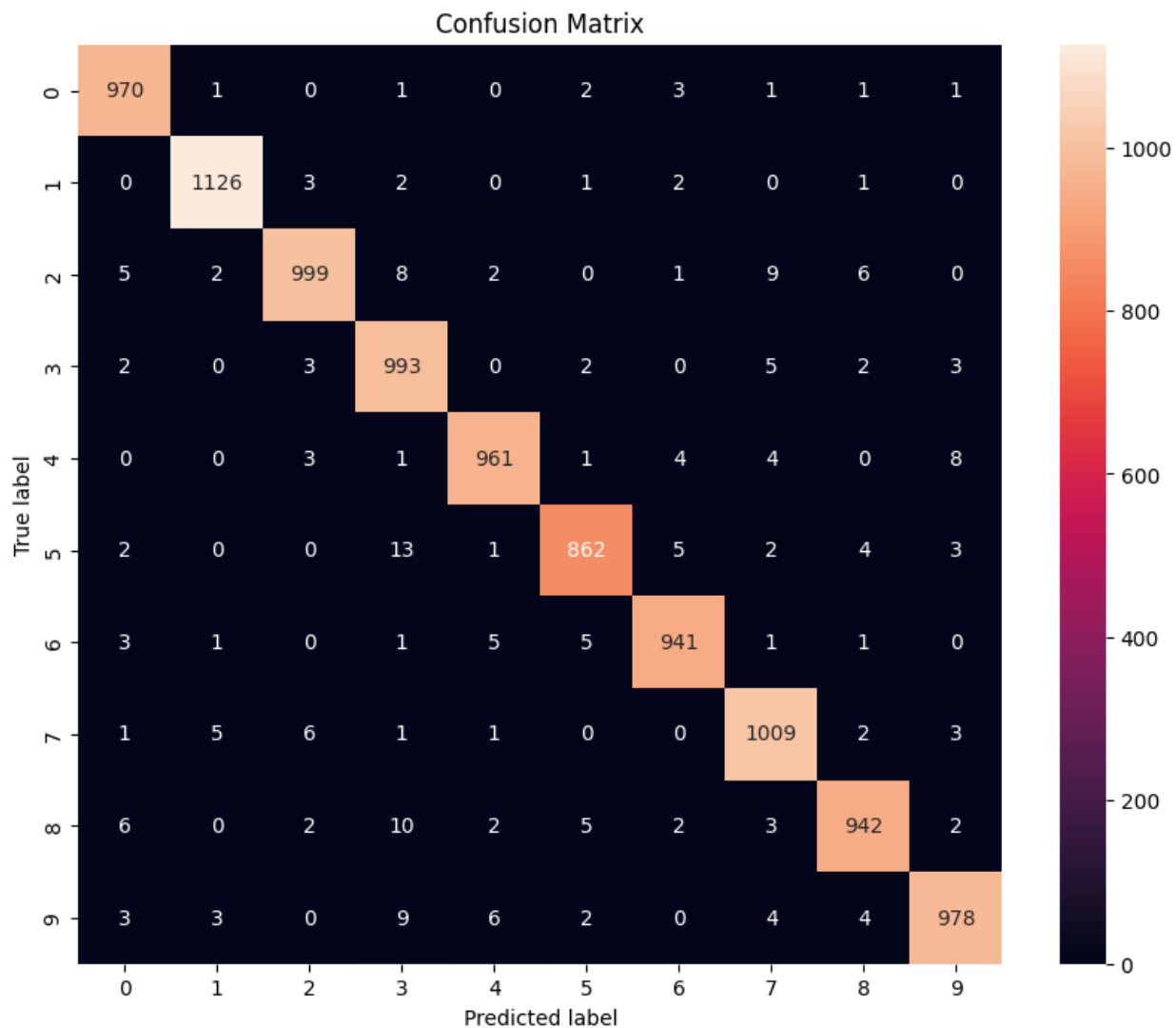
این شبکه به گونه‌ای طراحی شده است که تصاویر دست‌نویس اعداد را شناسایی و طبقه‌بندی کند. معماری ساده اما موثر آن برای مسائل پایه‌ای تشخیص الگو مانند MNIST مناسب است.



این نمودارها نتایج آموزش یک شبکه عصبی را نشان می‌دهند:

۱. نمودار سمت چپ (دقت آموزش و اعتبارسنجی): این نمودار دقت مدل را در طول زمان (از ۰ تا ۹ عصر) برای داده‌های آموزشی و اعتبارسنجی نشان می‌دهد. دقت آموزش (خط آبی) به طور پیوسته افزایش می‌یابد و به حدود ۹۹٪ می‌رسد، در حالی که دقت اعتبارسنجی (خط نارنجی) پس از افزایش اولیه، در حدود ۹۷٪ تثبیت می‌شود. این اختلاف ممکن است نشان‌دهنده بیش‌برازش (overfitting) باشد، جایی که مدل بر روی داده‌های آموزشی به خوبی عمل می‌کند اما در داده‌های جدید (اعتبارسنجی) نمی‌تواند همان عملکرد را نشان دهد.

۲. نمودار سمت راست (خطای آموزش و اعتبارسنجی): این نمودار خطای مدل را در طول زمان آموزش نشان می‌دهد. خطای آموزش (خط آبی) به طور مداوم کاهش می‌یابد و به نزدیکی ۰ میل می‌کند. از طرف دیگر، خطای اعتبارسنجی (خط نارنجی) پس از کاهش اولیه، تا حدودی ثابت می‌ماند و کاهش کمتری دارد. این نیز می‌تواند نشانه‌ای از بیش‌برازش باشد، زیرا مدل به بهبود عملکرد خود بر روی داده‌های آموزشی ادامه می‌دهد اما این بهبودها لزوماً در داده‌های اعتبارسنجی منعکس نمی‌شوند.



تصویر نشان داده شده ماتریس درهم‌ریختگی (Confusion Matrix) برای نتایج طبقه‌بندی دست‌نویس اعداد MNIST توسط یک مدل یادگیری ماشین است. ماتریس درهم‌ریختگی نحوه عملکرد یک مدل طبقه‌بندی را با نشان دادن تعداد پیش‌بینی‌های صحیح و نادرست برای هر کلاس نشان می‌دهد.

در این ماتریس، هر سطر نشان‌دهنده کلاس واقعی است و هر ستون نشان‌دهنده پیش‌بینی‌های مدل است. مقادیر روی قطر اصلی (از بالا چپ به پایین راست) نشان‌دهنده تعداد پیش‌بینی‌های صحیح برای هر کلاس است. برای مثال، مقدار ۹۷۰ نشان می‌دهد که ۹۷۰ مورد از اعداد '۰' به درستی توسط مدل شناسایی شده‌اند.



مقادیر خارج از قطر اصلی خطاها را نشان می‌دهند، مانند تعداد مواردی که یک عدد به اشتباه به عنوان عدد دیگری پیش‌بینی شده است.

بر اساس این ماتریس، می‌توان چندین نکته را استنباط کرد:

- مدل به طور کلی دقت بالایی دارد، زیرا تعداد زیادی از پیش‌بینی‌ها بر روی قطر اصلی هستند.
  - برخی از اعداد با دقت بیشتری نسبت به دیگران پیش‌بینی شده‌اند؛ برای مثال، عدد '۱' دارای بیشترین تعداد تشخیص‌های صحیح (۱۱۲۶ مورد) است.
  - خطاها نسبتاً کم هستند اما برخی الگوهای خطا وجود دارد؛ برای مثال، عدد '۵' اغلب به اشتباه به عنوان '۳' پیش‌بینی شده است (۱۳ مورد).
- دقت ۹۷٪ است.

	precision	recall	f1-score	support
0	0.98	0.99	0.98	980
1	0.99	0.99	0.99	1135
2	0.98	0.97	0.98	1032
3	0.96	0.98	0.97	1010
4	0.98	0.98	0.98	982
5	0.98	0.97	0.97	892
6	0.98	0.98	0.98	958
7	0.97	0.98	0.98	1028
8	0.98	0.97	0.97	974
9	0.98	0.97	0.97	1009
accuracy			0.98	10000
macro avg	0.98	0.98	0.98	10000
weighted avg	0.98	0.98	0.98	10000

بهترین زمان برای متوقف کردن آموزش شبکه زمانی است که دقت اعتبارسنجی دیگر بهبود نمی‌یابد یا شروع به کاهش می‌کند، در حالی که دقت آموزش همچنان در حال افزایش است. این می‌تواند نشانه‌ای از بیش‌برازش باشد، یعنی مدل در حال یادگیری جزئیات و نویز موجود در مجموعه داده آموزشی است به جای یادگیری الگوهای کلی که قابل تعمیم به داده‌های جدید هستند.

بر اساس نمودارهای دقت و خطای آموزش و اعتبارسنجی که قبلاً دیده‌ایم، می‌توان دید که در حدود اپوک ۴ تا ۵، دقت اعتبارسنجی تثبیت می‌شود و اختلاف بین دقت آموزش و اعتبارسنجی افزایش می‌یابد. این نقطه ممکن



است یک کاندیدای خوب برای توقف زودهنگام آموزش باشد. توقف زودهنگام یک تکنیک معمول برای جلوگیری از بیش‌برازش است که آموزش را متوقف می‌کند زمانی که عملکرد بر روی مجموعه داده اعتبارسنجی دیگر بهبود نمی‌یابد یا کاهش می‌یابد.

توجه به ماتریس درهم‌ریختگی نیز می‌تواند مفید باشد. اگر می‌بینیم که برخی از کلاس‌ها به طور مداوم اشتباه تشخیص داده می‌شوند، ممکن است بخواهیم آموزش را ادامه دهیم با استفاده از تکنیک‌هایی مانند تنظیم مجدد وزن یا افزایش داده‌های آموزشی برای بهبود تشخیص در آن کلاس‌ها.

**بخش ۲:** کد ارائه شده یک شبکه عصبی پیچشی (Convolutional Neural Network - CNN) را برای داده‌های MNIST ایجاد و آموزش می‌دهد. در زیر پارامترهای شبکه و توضیحات مربوط به تعداد و نوع لایه‌های مختلف آورده شده است:

۱. لایه‌های شبکه:

- `Conv2D(32, (3, 3), activation='relu')`: این اولین لایه پیچشی (convolutional) است که ۳۲ فیلتر پیچشی با اندازه  $3 \times 3$  دارد و از تابع فعال‌سازی 'relu' استفاده می‌کند. این لایه ویژگی‌های محلی تصاویر را استخراج می‌کند.

- `MaxPooling2D((2, 2))`: لایه‌ای است که از عملیات استخراج نمونه (pooling) با پنجره  $2 \times 2$  برای کاهش ابعاد ویژگی‌های استخراج شده استفاده می‌کند.

- `Conv2D(64, (3, 3), activation='relu')`: دومین لایه پیچشی با ۶۴ فیلتر پیچشی است که به استخراج ویژگی‌های پیچیده‌تر کمک می‌کند.

- `MaxPooling2D((2, 2))`: یک لایه استخراج نمونه دیگر که ابعاد ویژگی‌ها را بیشتر کاهش می‌دهد.

- `Conv2D(64, (3, 3), activation='relu')`: سومین و آخرین لایه پیچشی با ۶۴ فیلتر که ویژگی‌های بیشتری را فراهم می‌آورد.

- `Flatten()`: این لایه خروجی‌های استخراج شده از لایه‌های پیچشی و استخراج نمونه را به یک بردار یک‌بعدی تبدیل می‌کند تا بتوان آنها را به لایه‌های چگال (Dense) وصل کرد.



- Dense(128, activation='relu'): این یک لایه چگال با ۱۲۸ نورون است که تابع فعال‌سازی 'relu' دارد. این لایه می‌تواند ارتباطات پیچیده‌تری بین ویژگی‌های استخراج شده ایجاد کند.

- Dense(10, activation='softmax'): لایه خروجی شبکه است که ۱۰ نورون دارد، هر کدام نماینده یکی از کلاس‌های ده‌گانه اعداد (۰ تا ۹). تابع فعال‌سازی 'softmax' احتمالات پیش‌بینی شده برای هر کلاس را می‌دهد.

## ۲. پارامترهای مدل:

- optimizer='adam': بهینه‌ساز Adam است که نرخ یادگیری را به طور تطبیقی تنظیم می‌کند.

- loss='categorical\_crossentropy': تابع زیان برای مسائل دسته‌بندی چند کلاسی است.

- metrics=['accuracy']: معیار انتخابی برای ارزیابی مدل دقت طبقه‌بندی است.

## ۳. پیش‌پردازش داده‌ها:

- داده‌ها به مقادیر بین ۰ و ۱ نرمال‌سازی می‌شوند.

- داده‌های آموزشی به دو قسمت آموزش و اعتبارسنجی تقسیم می‌شوند.

- برچسب‌ها به فرمت one-hot encoding تبدیل می‌شوند.

## ۴. آموزش مدل:

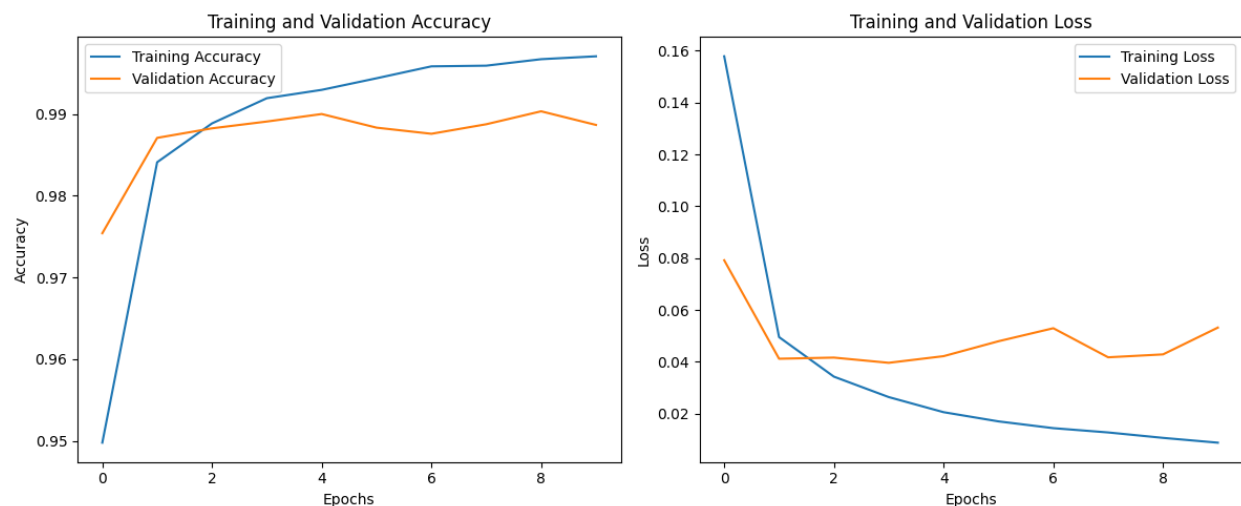
- مدل بر روی داده‌های آموزشی برای ۱۰ دوره (epoch) آموزش داده می‌شود و دقت و زیان بر روی داده‌های اعتبارسنجی ارزیابی می‌شود.

این معماری CNN نسبت به معماری چندلایه ساده (MLP) که قبلاً بررسی کردیم، پیچیده‌تر است و معمولاً در تشخیص تصویر عملکرد بهتری دارد، زیرا قادر است ویژگی‌های مکانی را در تصاویر استخراج کند. CNN‌ها به



خصوص در دسته‌بندی تصاویر که دارای الگوهای مکانی هستند، مانند دست‌نویس‌ها یا شناسایی چهره، بسیار موثرند.

با توجه به این که ورودی‌ها و برچسب‌ها به درستی پیش‌پردازش شده‌اند، مدل به خوبی قادر است از داده‌های آموزشی برای یادگیری و سپس تعمیم به داده‌های اعتبارسنجی و تست استفاده کند. این مدل به احتمال زیاد عملکرد بهتری نسبت به یک MLP ساده بر روی مجموعه داده MNIST از تصاویر دست‌نویس خواهد داشت.



این تصویر دو نمودار را نشان می‌دهد که عملکرد یک شبکه عصبی پیچشی (CNN) را در طول آموزش بر روی داده‌های MNIST تصویر می‌کنند.

نمودار سمت چپ (دقت آموزش و اعتبارسنجی):

نمودار دقت، افزایش مداوم دقت آموزش (خط آبی) را نشان می‌دهد که در نهایت به حدود ۹۹٪ می‌رسد.

دقت اعتبارسنجی (خط نارنجی) ابتدا به سرعت افزایش می‌یابد ولی پس از اپوک ۴، روند افزایش آهسته‌تر می‌شود و در حدود ۹۸.۵٪ ثابت می‌ماند.

نمودار سمت راست (خطای آموزش و اعتبارسنجی):

نمودار خطا، کاهش سریع خطای آموزش (خط آبی) را در ابتدای آموزش نشان می‌دهد که به نظر می‌رسد در اپوک‌های بعدی به طور پایدار ادامه می‌یابد.





خطای اعتبارسنجی (خط نارنجی) نیز به سرعت کاهش می‌یابد ولی پس از اپوک ۴، کاهش خطا کمتر شده و حتی نوساناتی را تجربه می‌کند.

تحلیل نمودارها نکات زیر را نشان می‌دهد:

دقت: مدل بر روی داده‌های آموزشی عملکرد بسیار خوبی دارد و دقت بالایی را نشان می‌دهد. دقت بر روی داده‌های اعتبارسنجی نیز بالا است، اما اختلاف نسبی بین دقت آموزش و اعتبارسنجی ممکن است نشان‌دهنده آغاز بیش‌برازش باشد،

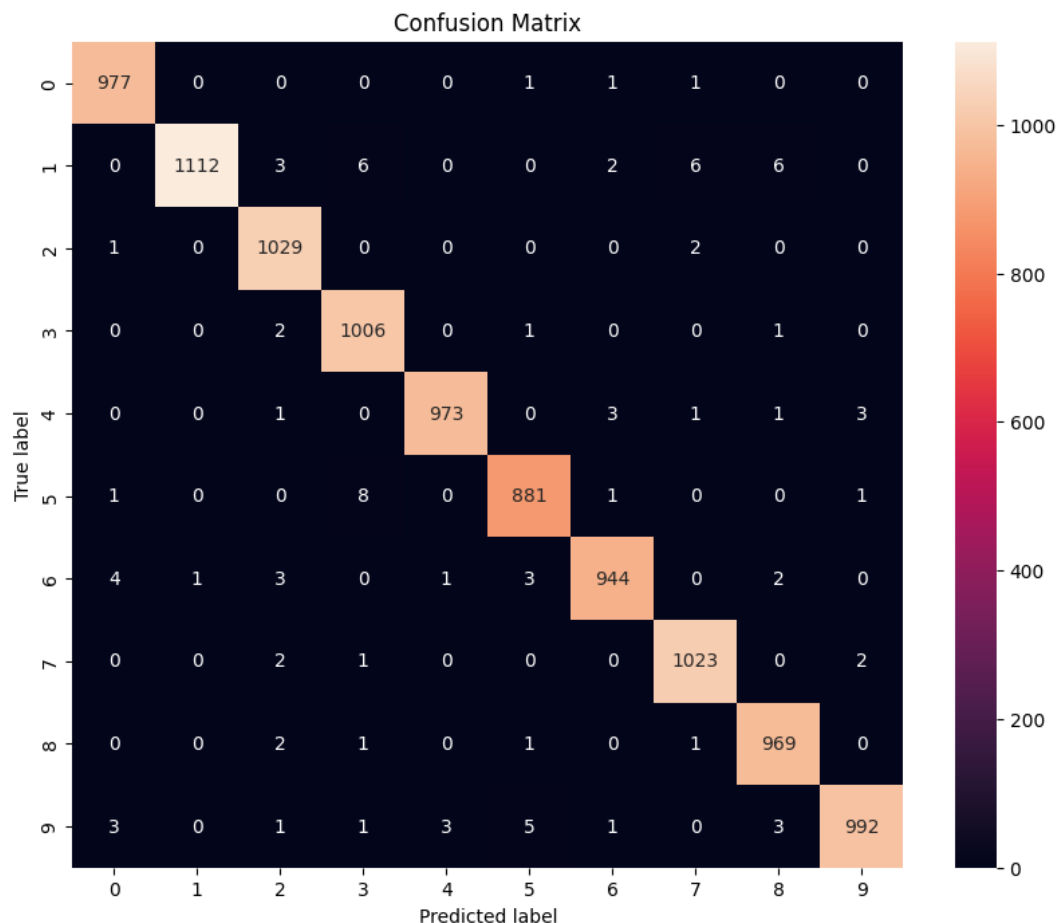
به ویژه اگر این اختلاف شروع به افزایش کند. با این حال، تا اپوک ۹، اختلاف زیادی بین دقت آموزش و اعتبارسنجی دیده نمی‌شود، که نشان‌دهنده این است که مدل نسبتاً خوبی در مقابله با بیش‌برازش عمل کرده است.

خطا: خطای آموزش و اعتبارسنجی هر دو کاهش یافته‌اند، که نشان‌دهنده یادگیری موثر مدل است. با این حال، نوسان در خطای اعتبارسنجی پس از اپوک ۴ نشان می‌دهد که مدل ممکن است دیگر بهبود چشمگیری در داده‌های اعتبارسنجی نداشته باشد.

بر اساس این تحلیل، می‌توان گفت که بهترین زمان برای متوقف کردن آموزش احتمالاً در اطراف اپوک ۵ یا ۶ است، زمانی که خطای اعتبارسنجی دیگر بهبود قابل توجهی ندارد و دقت نیز ثابت مانده است.

```
313/313 [=====] - 1s 4ms/step - loss: 0.0415 - accuracy: 0.9906
```

دقت طبقه بندی ۹۶٪ است.



از ماتریس درهم‌ریختگی می‌توان موارد زیر را استنباط کرد:

- مدل به طور کلی عملکرد بسیار خوبی دارد، زیرا بیشترین تعدادها در قطر اصلی قرار دارند که نشان‌دهنده پیش‌بینی‌های صحیح است.

- عدد '۱' دارای بیشترین تعداد تشخیص‌های صحیح (۱۱۱۲ مورد) است و کمترین خطا را دارد.

- برخی از اعداد مانند '۵' خطاهای بیشتری دارند؛ مثلاً '۵' در ۸ مورد به اشتباه به عنوان '۳' تشخیص داده شده است.

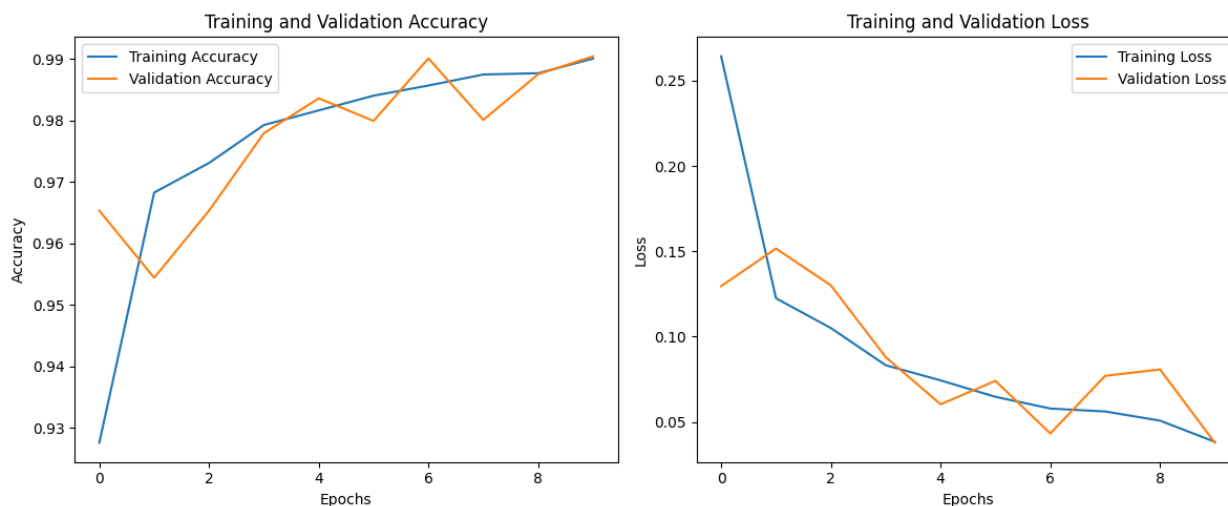
- کمترین تعداد تشخیص‌های صحیح برای عدد '۵' است که نشان‌دهنده ممکن است مدل در تشخیص این عدد با دشواری روبرو باشد.

این ماتریس می‌تواند به تحلیل‌گران کمک کند تا نقاط ضعف مدل را شناسایی کنند و رویکردهایی برای بهبود تشخیص در آن کلاس‌هایی که دارای بیشترین خطا هستند اتخاذ کنند. به عنوان مثال، ممکن است لازم باشد



داده‌های بیشتری برای کلاس‌هایی با خطای بالا جمع‌آوری کنیم، یا ممکن است تغییراتی در معماری مدل ایجاد کنیم تا ویژگی‌های مربوط به آن اعداد بهتر استخراج شوند. همچنین، اعمال تکنیک‌هایی مانند Data Augmentation یا تنظیم وزن‌ها (Regularization) می‌تواند به کاهش خطا کمک کند.

### بخش سه :



#### ۱. \*\*نمودار دقت (سمت چپ)\*\*:

- دقت آموزش (خط آبی) و دقت اعتبارسنجی (خط نارنجی) هر دو به طور کلی افزایشی هستند، که نشان‌دهنده بهبود عملکرد مدل با گذشت زمان است.

- دقت آموزش به‌طور مداوم افزایش یافته و در پایان به بیش از ۹۸٪ رسیده است.

- دقت اعتبارسنجی نوسانات بیشتری دارد اما در کل نیز افزایشی است. نوسانات دقت اعتبارسنجی می‌تواند ناشی از اندازه نمونه کوچک‌تر داده‌های اعتبارسنجی یا ویژگی‌های خاص داده‌ها باشد.

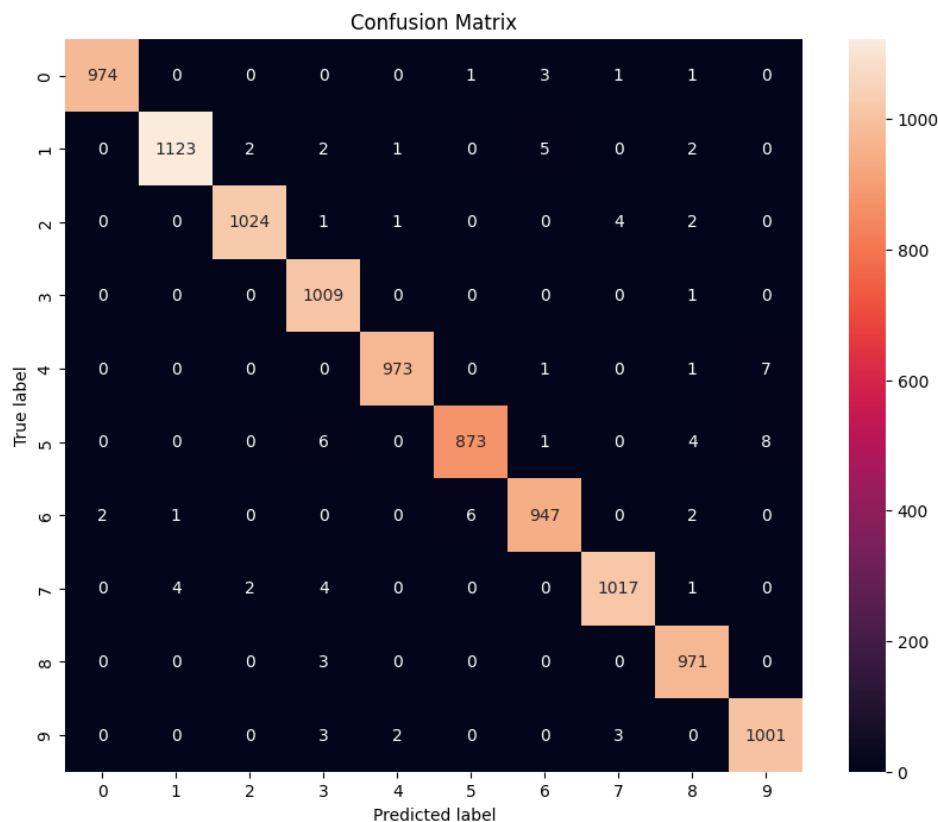
#### ۲. \*\*نمودار خطا (سمت راست)\*\*:

- خطای آموزش (خط آبی) به شدت در ابتدای آموزش کاهش یافته و پس از آن به طور تدریجی کاهش پیدا می‌کند.

- خطای اعتبارسنجی (خط نارنجی) نیز کاهش می‌یابد، اما نوسان بیشتری دارد. این نوسان می‌تواند نشان‌دهنده این باشد که مدل به داده‌های جدید اعتبارسنجی واکنش نشان می‌دهد و در هر اپوک یادگیری متفاوتی دارد.

با توجه به این داده‌ها، می‌توان چند نکته کلیدی را استنباط کرد:

- مدل شبکه MobileNet در حال پیشرفت است و با گذشت هر اپوک، بهتر عمل می‌کند.
- عدم وجود فاصله زیاد بین دقت آموزش و اعتبارسنجی نشان می‌دهد که مدل دچار بیش‌برازش (overfitting) قابل توجهی نشده است، اگرچه نوسانات در دقت اعتبارسنجی ممکن است نشان‌دهنده شروع بیش‌برازش باشد. این موضوع می‌تواند با تنظیم مجدد مدل (مانند افزایش داده‌های آموزشی، استفاده از تکنیک‌های منظم‌سازی یا dropout) برطرف شود.
- نوسان در خطای اعتبارسنجی ممکن است نشان‌دهنده این باشد که مدل ممکن است از یک نقطه خاصی به بعد به سمت بیش‌برازش پیش رود. اگر این نوسانات ادامه‌دار باشد و گسترش یابد، ممکن است نیاز به توقف زودهنگام یا تغییر در پارامترهای آموزش باشد.
- به طور کلی، مدل نشانه‌های خوبی از یادگیری و تعمیم‌پذیری نشان می‌دهد. برای تصمیم‌گیری دقیق در مورد اینکه آیا باید آموزش را ادامه دهیم یا خیر، باید علاوه بر نمودارها، به دیگر معیارهای عملکرد و همچنین دانش موضوعی در مورد داده‌ها و کاربرد خاص مدل توجه داشت.





بر اساس تصویر می‌توان موارد زیر را تحلیل کرد:

- عموماً، مدل عملکرد خوبی داشته است چرا که بیشترین تعداد پیش‌بینی‌های صحیح در قطر اصلی قرار دارند، نشان‌دهنده دقت بالای مدل در طبقه‌بندی اعداد.

- برای بیشتر کلاس‌ها، تعداد خیلی کمی از تصاویر به اشتباه طبقه‌بندی شده‌اند. به عنوان مثال، تنها یک تصویر از عدد '۰' به عنوان '۴' و سه تصویر به عنوان '۲' طبقه‌بندی شده است.

- برخی از کلاس‌ها مانند '۵' و '۷' به نظر می‌رسد که بیشتر از بقیه دچار اشتباه در طبقه‌بندی شده‌اند. به عنوان مثال، '۵' به اشتباه ۶ بار به عنوان '۳' و ۸ بار به عنوان '۸' طبقه‌بندی شده است.

- عدد '۷' نیز دارای چندین مورد اشتباه طبقه‌بندی است، که ۴ بار به عنوان '۲'، ۲ بار به عنوان '۳'، و ۴ بار به عنوان '۹' طبقه‌بندی شده است.

- این اطلاعات می‌توانند برای بهبود مدل استفاده شوند، به طور مثال از طریق جمع‌آوری یا تولید نمونه‌های بیشتر از کلاس‌هایی که دارای خطای بالایی هستند، یا اصلاح معماری شبکه برای حساس‌سازی بیشتر به ویژگی‌های منحصر به فرد این اعداد.

نکته قابل توجه دیگر این است که خطاهای طبقه‌بندی در برخی موارد ممکن است به دلیل شباهت‌های ظاهری بین برخی از اعداد باشد. به عنوان مثال، اشتباه بین '۵' و '۳' یا بین '۷' و '۱' ممکن است به خاطر شباهت‌هایی در نحوه نوشتار این اعداد باشد.

پاسخ ۶.

روش انجام:

محاسبه Entropy و Information Gain: برای هر ویژگی، باید انتروپی و سود اطلاعاتی (Information Gain) را محاسبه کنید تا بهترین ویژگی برای تقسیم هر گره را پیدا کنید.

ساخت درخت تصمیم: با استفاده از ویژگی با بیشترین سود اطلاعاتی، داده‌ها را تقسیم کنید و برای هر زیرمجموعه، گره‌های جدید درخت را ایجاد کنید. این فرآیند را تکرار کنید تا همه داده‌ها در یک گره به یک کلاس تعلق داشته باشند یا هیچ ویژگی دیگری برای تقسیم باقی نمانده باشد.



برچسب‌زنی گره‌های برگ: برای گره‌هایی که نمی‌توانند بیشتر تقسیم شوند، یک برچسب کلاس بر اساس بیشترین تعداد نمونه‌های موجود در آن گره اختصاص دهید.

پیش‌بینی و ارزیابی: با استفاده از درخت تصمیم ساخته شده، برای هر نمونه در داده‌های آزمایش، پیش‌بینی انجام دهید و دقت مدل را محاسبه کنیم

مراحل در کد پیوست شده موجود است.

```
Train Accuracy: 0.8731  
Test Accuracy: 0.8731
```

مقدار دقت مدل به شرح بالا می‌باشد