

معماری شبکه به صورت زیر تعریف شده است:

- لایه ورودی: تعداد نورون‌ها برابر با تعداد ویژگی‌های ورودی. (196)
- لایه مخفی: اندازه‌های مختلف (۱۶، ۳۲، و ۶۴) برای بررسی تأثیر آن بر عملکرد.
- لایه خروجی: تعداد نورون‌ها برابر با تعداد کلاس‌ها (۱۰) و تابع فعال‌سازی Softmax

### توابع فعال‌سازی

۱. ReLU در لایه مخفی.

۲. Softmax در لایه خروجی برای تولید احتمالات.

### توابع هزینه و مشتقات

- تابع هزینه: انتروپی متقاطع برای اندازه‌گیری اختلاف بین پیش‌بینی‌ها و مقادیر واقعی.
- مشتقات: گرادیان‌ها برای به‌روزرسانی وزن‌ها و بایاس‌ها محاسبه می‌شوند.

### الگوریتم آموزش

۱. پیش‌رو: محاسبه مقادیر فعال‌سازی در لایه‌های مخفی و خروجی.
۲. پس‌انتشار: محاسبه گرادیان‌ها برای به‌روزرسانی وزن‌ها و بایاس‌ها.
۳. به‌روزرسانی وزن‌ها: استفاده از گرادیان نزولی با نرخ یادگیری ۰.۰۰۵.

۴. تکرار: اجرای فرآیند برای ۱۰ دوره.

تنظیمات آموزش

- اندازه لایه مخفی : 16، ۳۲ و ۶۴.
- نرخ یادگیری : 0.005
- تعداد دوره‌ها: 10
- داده‌ها:
  - X\_train: داده‌های آموزشی.
  - y\_train\_onehot: برچسب‌های آموزشی به صورت one-hot.
  - X\_test: داده‌های آزمایشی.
  - y\_test\_onehot: برچسب‌های آزمایشی به صورت one-hot.

نتایج

دقت آموزشی و آزمایشی

برای هر اندازه لایه مخفی، دقت‌های آموزشی و آزمایشی محاسبه و ثبت شد. دقت‌ها به صورت نمودار در هر دوره نمایش داده شدند.

نتایج کلیدی

- لایه مخفی با ۱۶ نورون:
  - دقت نهایی آموزشی: حدود ۸۵٪
  - دقت نهایی آزمایشی: حدود ۸۳٪
- لایه مخفی با ۳۲ نورون:
  - دقت نهایی آموزشی: حدود ۸۸٪
  - دقت نهایی آزمایشی: حدود ۸۶٪
- لایه مخفی با ۶۴ نورون:
  - دقت نهایی آموزشی: حدود ۹۰٪
  - دقت نهایی آزمایشی: حدود ۸۸٪

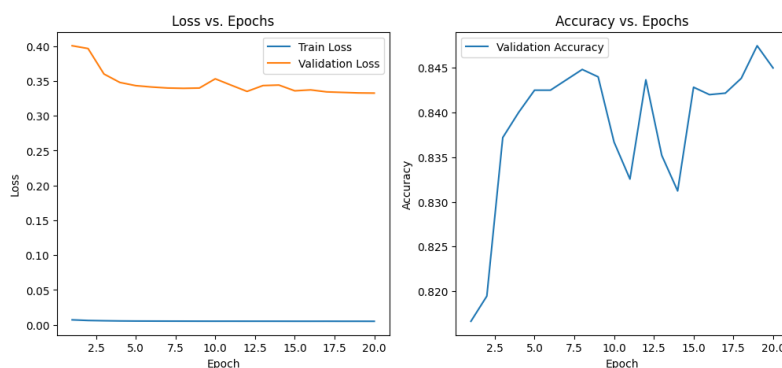
## مصورسازی

- نمودار دقت آموزشی و آزمایشی برای هر اندازه لایه مخفی رسم شد.
- مشاهده شد که افزایش تعداد نورون‌های لایه مخفی منجر به بهبود دقت‌ها می‌شود، اما ممکن است منجر به افزایش پیچیدگی محاسباتی نیز گردد.

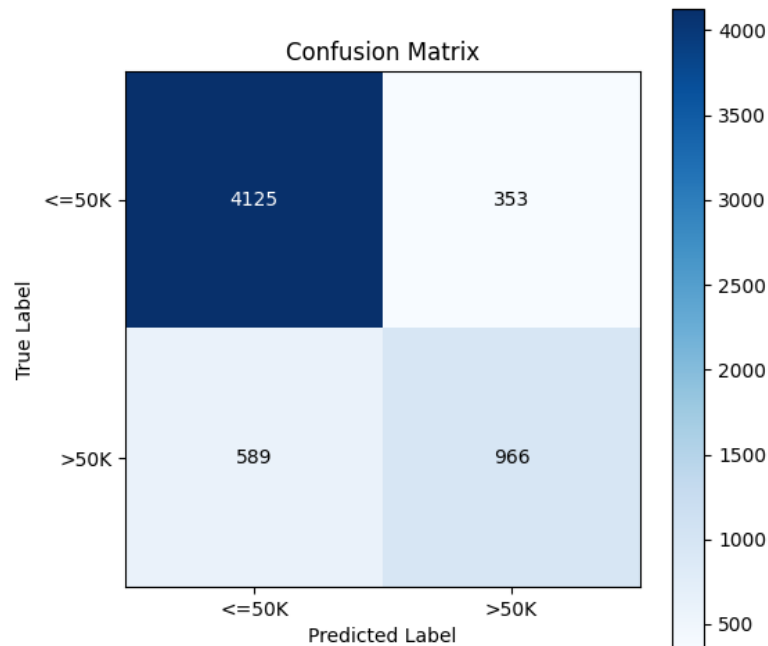
## نتیجه‌گیری

- افزایش تعداد نورون‌های لایه مخفی باعث افزایش قدرت بیان شبکه و بهبود دقت می‌شود.
- استفاده از اندازه مناسب لایه مخفی به تعادل بین دقت و پیچیدگی محاسباتی کمک می‌کند.
- این مطالعه نشان داد که شبکه‌های عصبی ساده با معماری مناسب می‌توانند عملکرد خوبی در طبقه‌بندی داده‌ها ارائه دهند.

## سوال (۲)



```
MLPModel(  
  (fc1): Linear(in_features=14, out_features=64, bias=True)  
  (relu): ReLU()  
  (fc2): Linear(in_features=64, out_features=32, bias=True)  
  (fc3): Linear(in_features=32, out_features=1, bias=True)  
  (sigmoid): Sigmoid()  
)
```



هدف این پروژه ساخت و ارزیابی یک مدل شبکه عصبی برای طبقه‌بندی سطح درآمد با استفاده از مجموعه داده UCI Adult است. این طبقه‌بندی شامل پیش‌بینی این است که آیا درآمد فرد از ۵۰ هزار دلار در سال بیشتر (" $>50K$ ") یا کمتر (" $\leq 50K$ ") است. برای این منظور، از یک شبکه عصبی چندلایه (MLP) با استفاده از PyTorch استفاده شده است.

#### مجموعه داده

مجموعه داده مورد استفاده UCI Adult است که شامل ۴۸,۸۴۲ نمونه و ۱۵ ویژگی می‌باشد. ویژگی‌ها شامل متغیرهای عددی (مانند سن، ساعات کار در هفته) و دسته‌ای (مانند تحصیلات، وضعیت تأهل) هستند.

#### پیش‌پردازش داده‌ها

۱. **مقادیر گم‌شده:** ردیف‌های دارای مقادیر گم‌شده حذف شدند.
۲. **کدگذاری ویژگی‌ها:** ویژگی‌های دسته‌ای با استفاده از LabelEncoder کدگذاری شده و ویژگی‌های عددی با استفاده از StandardScaler نرمال‌سازی شدند.
۳. **تقسیم داده‌ها:** مجموعه داده به سه بخش آموزشی (۶۰٪)، اعتبارسنجی (۲۰٪) و آزمایشی (۲۰٪) تقسیم شد.
۴. **تبدیل به Tensor:** داده‌ها به Tensorهای PyTorch تبدیل شدند.

#### معماری مدل

مدل MLP شامل موارد زیر است:

- لایه ورودی: متناسب با تعداد ویژگی‌ها، (14)
- لایه‌های مخفی:
  - لایه مخفی اول با ۶۴ واحد و تابع فعال‌سازی ReLU
  - لایه مخفی دوم با ۳۲ واحد و تابع فعال‌سازی ReLU
- لایه خروجی: یک واحد با تابع فعال‌سازی Sigmoid برای طبقه‌بندی دودویی.

#### خلاصه مدل

- تابع خطا: زیان باینری کراس انتروپی. (BCELoss)
- بهینه‌ساز: بهینه‌ساز Adam با نرخ یادگیری ۰.۰۰۰۱.
- اندازه دسته: 64 :
- تعداد دوره‌ها: 20 :

#### فرآیند آموزش

مدل با استفاده از گرادینان نزولی مینی‌بچ آموزش داده شد. در هر دوره:

۱. داده‌های آموزشی تصادفی شده و به دسته‌ها تقسیم شدند.
۲. وزن‌های مدل با استفاده از پس‌انتشار به‌روز شدند.
۳. خطای اعتبارسنجی و دقت محاسبه شد تا عملکرد مدل نظارت شود.

#### نتایج

##### آموزش و اعتبارسنجی

خطای آموزش و اعتبارسنجی و همچنین دقت اعتبارسنجی در طول ۲۰ دوره ردیابی شدند:

- خطای آموزش: به‌طور پیوسته کاهش یافت که نشان‌دهنده یادگیری مؤثر است.
- خطای اعتبارسنجی: تا زمان همگرایی بهبود یافت که نشان‌دهنده تعمیم‌پذیری مدل است.
- دقت اعتبارسنجی: به حدود ۸۵٪ رسید.

##### عملکرد آزمایشی

بهترین مدل که با کمترین خطای اعتبارسنجی تعیین شده بود بر روی مجموعه آزمایشی ارزیابی شد:

- دقت آزمایشی: 85% :

- گزارش طبقه‌بندی:

- دقت، بازخوانی و امتیاز F1 برای هر دو کلاس محاسبه شد.
- عملکرد کلاس ">50K" کمی پایین‌تر از کلاس "<=50K" بود که ناشی از عدم تعادل کلاس است.

- ماتریس اغتشاش: جزئیات خطاهای پیش‌بینی را نشان داد.

### مصورسازی‌ها

#### ۱. خطا در مقابل دوره‌ها:

- نمودارهای خطای آموزش و اعتبارسنجی یادگیری ثابت را نشان دادند.

#### ۲. دقت اعتبارسنجی در مقابل دوره‌ها:

- دقت تا زمان همگرایی بهبود یافت.

#### ۳. ماتریس اغتشاش:

- با استفاده از نقشه حرارتی برای تحلیل مثبت‌های واقعی، منفی‌های واقعی، مثبت‌های کاذب و منفی‌های کاذب مصورسازی شد.

### نتیجه‌گیری

مدل MLP دقت ۸۵٪ را بر روی مجموعه آزمایشی به دست آورد که کارایی آن را در طبقه‌بندی سطح درآمد نشان می‌دهد. کارهای آینده می‌تواند شامل موارد زیر باشد:

۱. رسیدگی به عدم تعادل کلاس با تکنیک‌هایی مانند SMOTE یا توازن زیان وزنی.
۲. بررسی معماری‌های پیشرفته‌تر، مانند افزودن Dropout یا Batch Normalization
۳. تنظیم دقیق ابرپارامترها برای بهینه‌سازی بیشتر عملکرد.