



تمرین سری دوم  
درس یادگیری عمیق

نام مدرس: دکتر محمدرضا محمدی

دستیاران آموزشی مرتبط: دنیا اسفندیارپور،  
ساره افتخاری، حجت شهبابی، نازنین احمدپور

مهلت تحویل (بدون کسر نمره):  
۲۴ آبان ماه

سوالات تئوری (۶۰ نمره)

سوال ۱ (۷ نمره)

یک تیم تحقیقاتی قصد دارد یک مدل شبکه عصبی عمیق برای پیش‌بینی قیمت مسکن طراحی کند. آن‌ها متوجه شده‌اند که مدلشان روی داده‌های آموزشی عملکرد بسیار خوبی دارد، اما روی داده‌های اعتبارسنجی (Validation) دچار افت شدید عملکرد است. در بررسی بیشتر مشخص شده:

- تعداد ویژگی‌های ورودی نسبتاً زیاد است.
  - برخی ویژگی‌ها همبستگی بالایی با یکدیگر دارند.
  - داده‌ها نویز دارند و نمونه‌ها نامتوازن هستند.
  - هنگام آموزش، مقدار وزن‌ها به مرور خیلی بزرگ می‌شوند و نوسان در تابع هزینه دیده می‌شود.
- این تیم تصمیم گرفته برای کنترل بیش‌برازش (Overfitting) و پایدارسازی آموزش، از روش‌های رگولاریزیشن استفاده کند. شما به عنوان مشاور تیم باید:
- الف) توضیح دهید کدام روش‌های رگولاریزیشن (حداقل ۳ مورد) برای این مسئله بیشتر مناسب هستند و چرا.
- ب) خطرات یا معایب احتمالی هر یک از روش‌های پیشنهادی را بیان کنید.
- پ) اگر تیم تصمیم بگیرد از Early Stopping یا Dropout استفاده کند، این روش‌ها در چه شرایطی می‌توانند نامناسب یا حتی مضر باشند؟ مثال بزنید.
- ت) با توجه به سناریو، آیا ترکیب چند روش رگولاریزیشن توصیه می‌شود؟ اگر بله، یک استراتژی ترکیبی معقول پیشنهاد دهید.

سوال دو (۱۰ نمره)

یک شرکت امنیتی در حال توسعه یک سیستم تشخیص تقلب مالی (Fraud Detection) با استفاده از شبکه‌های عصبی عمیق است. این سیستم باید تراکنش‌های بانکی را در لحظه تحلیل کند. ویژگی‌های داده و مدل به صورت زیر است:

- داده‌ها بسیار نامتوازن هستند (کمتر از ۰,۵٪ تراکنش‌ها تقلبی‌اند).
- مدل شبکه عصبی عمیق به دلیل نیاز به دقت بالا، بسیار پیچیده طراحی شده (تعداد پارامتر بالا).

- روی داده‌های آموزش عملکرد خوبی دارد اما به شدت دچار Overfitting شده.
- آزمایش‌های اولیه نشان می‌دهند که رگولاریزیشن L2 و Dropout به تنهایی کافی نیستند.
- تیم توسعه نگران از دست رفتن نمونه‌های مهم تقلبی در اثر رگولاریزیشن شدید هستند.

الف) چرا استفاده از تکنیک‌های کلاسیک رگولاریزیشن مثل L2 یا Dropout به تنهایی کافی نیستند و ممکن است حتی روی این مسئله اثر منفی بگذارند؟

ب) سه مکانیزم رگولاریزیشن که مخصوص داده‌های نامتوازن یا مسائل حساس مانند امنیت و تقلب هستند را جستجو و رفرنس‌دهی کرده و توضیح دهید چگونه مانع Overfitting می‌شوند.

پ) یکی از اعضای تیم پیشنهاد داده به جای رگولاریزیشن، ظرفیت مدل را کاهش دهیم یا لایه‌ها را کم کنیم. آیا این استراتژی منطقی است؟ تحلیل کنید.

ت) اگر قرار باشد از رگولاریزیشن مبتنی بر داده (Data-Level Regularization) استفاده شود، کدامین یک از روش‌های موجود را پیشنهاد می‌دهید و چگونه این روش‌ها خطر Underfitting را کاهش می‌دهند؟

### سوال سه (۱۵ نمره)

یک شبکه‌ی عصبی کوچک برای تشخیص بیماری از داده‌های آزمایشگاهی توسعه داده‌اید، مشکلی که وجود دارد این است که شبکه‌ی شما روی داده‌های آموزشی عالی کار می‌کند (دقت ۹۰٪) اما روی داده‌های تست فقط ۶۸٪ دقت دارد! برداشت شما این است که مدل الگوها را حفظ کرده و می‌خواهید loss function را تغییر دهید تا مدل الگوهای کلی‌تر یاد بگیرد.

شما تصمیم گرفتید با یک نمونه به صورت کاملاً دستی محاسبات را انجام دهید تا ببینید دقیقاً چه اتفاقی می‌افتد.

#### معماری شبکه:

- ورودی: ۳ ویژگی (نتایج آزمایش خون)

- لایه مخفی: ۴ نورون (ReLU)

- خروجی: ۲ نورون برای سالم / بیمار (softmax)

تعداد پارامترها:

- لایه اول:  $4 \times 3 = 12$  وزن + ۴ بایاس

- لایه دوم:  $2 \times 4 = 8$  وزن + ۲ بایاس

**W1:**

	$h_1$	$h_2$	$h_3$	$h_4$
$x_1$	[ 2.5	-1.2	0.8	3.1 ]
$x_2$	[-0.5	2.8	-1.5	0.3 ]
$x_3$	[ 1.7	-0.9	2.3	-1.8 ]

$\text{bias}_1 = [0.5, -0.3, 0.2, 0.7]$

**W2:**

	$y_1$	$y_2$
$h_1$	[ 1.5	-2.3 ]
$h_2$	[-0.8	1.9 ]
$h_3$	[ 2.1	-0.6 ]
$h_4$	[-1.3	2.7 ]

$\text{bias}_2 = [0.3, -0.5]$

$x = [1.0, 0.5, 2.0]$

$y_{\text{true}} = [1, 0]$  (کلاس سالم)

الف) درک وضعیت فعلی

- خروجی لایه مخفی را محاسبه کنید.
- خروجی نهایی را محاسبه کنید.
- Loss را حساب کنید.

$$\text{Loss} = -\sum y_i \log(\hat{y}_i)$$

ب) کنترل وزن ها

ایده A:  $L_{\text{new}} = L_{\text{data}} + 0.01 \sum w^2$

ایده B:  $L_{\text{new}} = L_{\text{data}} + 0.005 \sum |w|$

- مقدار کل جریمه را برای شبکه در هر دو ایده محاسبه کنید.
- Loss کل را با هر دو ایده محاسبه کنید.
- اگر گرادینان نسبت به  $W2[0,0]$  از روی داده برابر  $0.3$  - باشد، گرادینان کل را برای هر دو ایده محاسبه کنید.
- کدام ایده وزن را بیشتر تغییر می دهد؟ چرا؟
- کدام را انتخاب می کنید؟ چرا؟

## سوال چهار

شما به عنوان یک پژوهشگر در تیم بینایی ماشین یک شرکت پزشکی کار می کنید و هدف طراحی مدلی سبک و مقاوم در برابر overfitting برای تشخیص سلول های سرطانی از تصاویر میکروسکوپی (32x32 RGB) است. به دلیل محدودیت داده ها (فقط ۱۰۰۰ تصویر)، مدل شما به شدت مستعد overfitting است.

تیم تصمیم گرفته است که برای کنترل Overfitting از Dropout و انواع آن استفاده کند، شما مسئول بررسی دقیق این روش ها و پیشنهاد بهترین گزینه هستید.

### بخش ۱: درک مفهومی و ریاضی (۱۰ نمره)

(الف) توضیح دهید چرا Dropout پیشنهاد شد و چه نوع نویزی را وارد شبکه می کند.

(ب) از دید ریاضی نشان دهید چگونه Dropout در هنگام آموزش باعث تقریب ensemble مدل ها می شود.

(ج) بررسی کنید که چگونه اگر در یک لایه، خروجی نورون ها  $h=f(Wx+b)$  باشد، با dropout احتمال  $p$ ، خروجی جدید می شود:  $\tilde{h}=m \odot h$ ,  $m \sim \text{Bernoulli}(1-p)$

و درستی این نرمال سازی را با اثبات دستی بررسی کنید.

(د) مفهوم Inverted Dropout را دقیق بنویسید. تفاوت آن با dropout معمولی چیست؟ چرا در پیاده سازی های مدرن از آن استفاده می شود؟

(ه) به صورت تحلیلی نشان دهید که چگونه این نسخه میانگین خروجی را در زمان تست ثابت نگه می دارد.

### بخش ۲: بررسی Spatial Dropout و DropConnect (۱۰ نمره)

Spatial Dropout:

(الف) توضیح دهید چرا در شبکه های کانولوشنی از حذف تصادفی کل channel ها استفاده می شود نه پیکسل های منفرد.

(ب) فرمول ریاضی آن را بنویسید.

(ج) مزایا و معایب آن را نسبت به dropout معمولی بنویسید.

DropConnect:

الف) در این روش به جای حذف خروجی‌ها، وزن‌ها را به صورت تصادفی حذف کنید.

ب) فرمول ریاضی آن را بنویسید.

ج) نشان دهید از دید تئوری این روش چطور regularization قوی‌تری نسبت به dropout اعمال می‌کند.

بخش ۳: تحلیل مقایسه ای (۸ نمره)

جدول زیر را کامل کنید.

سپس تعیین کنید در این سوال کدام روش انتخاب می‌شود؟ چرا؟

روش	سطح اعمال	نوع نویز	مزیت	عیب	مناسب برای
dropout					
Inverted dropout					
dropconnect					
Spatial dropout					

سوالات عملی (۶۰ نمره)

لطفاً برای پاسخگویی به سوالات عملی به پوشه‌ی practical مراجعه کنید. (هر سوال ۱۵ نمره)

نکات تکمیلی:

دانشجویان محترم حتماً فایل قوانین را مطالعه کرده و در انجام و ارسال تمارین رعایت بفرمایید.

موفق و سربلند باشید.