



دانشگاه صنعتی امیرکبیر  
(پلی تکنیک تهران)  
دانشکده ...

پایان نامه کارشناسی ارشد  
گرایش ...

## عنوان پایان نامه - دستورالعمل و راهنمای نگارش پایان نامه

نگارش

نام و نام خانوادگی کامل نویسنده

استاد راهنما

نام کامل استاد راهنما

استاد مشاور

نام کامل استاد مشاور

ماه و سال

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

# صفحه فرم ارزیابی و تصویب پایان نامه - فرم تأیید اعضاء کمیته دفاع

در این صفحه فرم دفاع یا تأیید و تصویب پایان نامه موسوم به فرم کمیته دفاع - موجود در پرونده آموزشی - را قرار دهید.

## نکات مهم:

- نگارش پایان نامه/رساله باید به **زبان فارسی** و بر اساس آخرین نسخه دستورالعمل و راهنمای تدوین پایان نامه های دانشگاه صنعتی امیرکبیر باشد.(دستورالعمل و راهنمای حاضر)
- رنگ جلد پایان نامه/رساله چاپی کارشناسی، کارشناسی ارشد و دکترا باید به ترتیب مشکی، طوسی و سفید رنگ باشد.
- چاپ و صحافی پایان نامه/رساله بصورت **پشت و رو(دورو)** بلامانع است و انجام آن توصیه می شود.



دانشگاه صنعتی امیرکبیر  
(پلی تکنیک تهران)

به نام خدا

## تعهدنامه اصالت اثر

تاریخ: ماه و سال

اینجانب نام و نام خانوادگی کامل نویسنده متعهد می‌شوم که مطالب مندرج در این پایان‌نامه حاصل کار پژوهشی اینجانب تحت نظارت و راهنمایی اساتید دانشگاه صنعتی امیرکبیر بوده و به دستاوردهای دیگران که در این پژوهش از آنها استفاده شده است مطابق مقررات و روال متعارف ارجاع و در فهرست منابع و مآخذ ذکر گردیده است. این پایان‌نامه قبلاً برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نگردیده است.

در صورت اثبات تخلف در هر زمان، مدرک تحصیلی صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پیگیری قانونی خواهد داشت.

کلیه نتایج و حقوق حاصل از این پایان‌نامه متعلق به دانشگاه صنعتی امیرکبیر می‌باشد. هرگونه استفاده از نتایج علمی و عملی، واگذاری اطلاعات به دیگران یا چاپ و تکثیر، نسخه‌برداری، ترجمه و اقتباس از این پایان‌نامه بدون موافقت کتبی دانشگاه صنعتی امیرکبیر ممنوع است. نقل مطالب با ذکر مآخذ بلامانع است.

نام و نام خانوادگی کامل نویسنده

امضا

نویسنده پایان نامه، در صورت تمایل میتواند برای پاسخگویی پایان نامه خود را به شخص یا اشخاص و یا ارگان خاصی تقدیم نماید.

## پاس کزاری

نویسنده پایان نامه می تواند مراتب امتنان خود را نسبت به استاد راهنما و استاد مشاور و یا دیگر افرادی که طی انجام پایان نامه به نحوی او را یاری و یا با او همکاری نموده اند ابراز دارد.

نام و نام خانوادگی کامل نویسنده

ماه و سال

## چکیده

در این قسمت چکیده پایان نامه نوشته می‌شود. چکیده باید جامع و بیان‌کننده خلاصه‌ای از اقدامات انجام‌شده باشد. در چکیده باید از ارجاع به مرجع و ذکر روابط ریاضی، بیان تاریخچه و تعریف مسئله خودداری شود.

## واژه‌های کلیدی:

کلیدواژه اول، ...، کلیدواژه پنجم (نوشتن سه تا پنج واژه کلیدی ضروری است)

# فهرست مطالب

عنوان

صفحه

۱	راهنمای استفاده از الگوی لاتک دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران) . . . .	۱
۱-۱	مقدمه . . . . .	۲
۲-۱	این همه فایل؟! . . . . .	۲
۳-۱	از کجا شروع کنم؟ . . . . .	۳
۴-۱	مطالب پایان نامه را چگونه بنویسم؟ . . . . .	۴
۱-۴-۱	نوشتن فصل ها . . . . .	۴
۲-۴-۱	مراجع . . . . .	۵
۳-۴-۱	واژه نامه فارسی به انگلیسی و برعکس . . . . .	۵
۵-۱	اگر سوالی داشتیم، از کی بپرسیم؟ . . . . .	۵
۲	مرور کارهای پیشین . . . . .	۶
۱-۲	یادگیری پیوسته . . . . .	۷
۱-۱-۲	چالش ها . . . . .	۷
۲-۱-۲	سناریوها . . . . .	۷
۳-۱-۲	رویکردها . . . . .	۸
۴-۱-۲	کاربردها . . . . .	۱۰
۲-۲	یادگیری پیوسته در بینایی کامپیوتر . . . . .	۱۱
۱-۲-۲	دسته بندی تصویر . . . . .	۱۱
۲-۲-۲	تشخیص عمل . . . . .	۱۳
۳-۲	مدل های بینایی-زبان . . . . .	۱۵
۱-۳-۲	پیش آموزش مدل های بینایی-زبان . . . . .	۱۶
۲-۳-۲	یادگیری انتقالی مدل های بینایی-زبان . . . . .	۱۷
۳-۳-۲	تقطیر دانش . . . . .	۱۸
۴-۲	یادگیری پیوسته در مدل های بینایی-زبان . . . . .	۱۹
۱-۴-۲	روش های مبتنی بر حافظه . . . . .	۲۰
۲-۴-۲	روش های مبتنی بر تنظیم . . . . .	۲۰



۲۰	۳-۴-۲ روش‌های مبتنی بر تقطیر دانش
۲۱	۴-۴-۲ روش‌های مبتنی بر معماری
۲۶	۳ نگارش صحیح
۲۷	۱-۳ مقدمه
۲۷	۲-۳ فارسی نویسی
۲۸	۳-۳ رعایت املای صحیح
۲۸	۴-۳ رعایت قواعد نشانه‌گذاری
۲۸	۱-۴-۳ ویرگول
۲۹	۲-۴-۳ نقطه
۲۹	۳-۴-۳ دونقطه
۲۹	۴-۴-۳ گیومه
۲۹	۵-۴-۳ نشانه پرشی
۳۰	۶-۴-۳ خط تیره
۳۰	۷-۴-۳ پرانتز
۳۰	۵-۳ جدا یا سرهم نوشتن برخی کلمات
۳۲	۴ مشخصات یک پایان نامه و گزارش علمی
۳۳	۱-۴ برخورداری از غنای علمی
۳۳	۲-۴ ارجاع به‌موقع و صحیح به منابع دیگر
۳۳	۳-۴ ساده‌نویسی
۳۴	۴-۴ وحدت موضوع
۳۴	۵-۴ اختصار
۳۴	۶-۴ رعایت نکات دستوری و نشانه‌گذاری
۳۴	۷-۴ توجه به معلومات ذهنی مخاطب
۳۴	۸-۴ رعایت مراحل اصولی نگارش
۳۶	۵ جمع‌بندی و نتیجه‌گیری و پیشنهادات
۳۷	۱-۵ پیشنهادات

۳۸	۶ موارد به روزرسانی
۳۹	۱-۶ استفاده از subfigure
۳۹	۲-۶ فرض
۴۰	۳-۶ جدول
۴۰	۱-۳-۶ چندسطری
۴۰	۲-۳-۶ مقیاس بندی
۴۱	۳-۳-۶ استفاده از قضیه
۴۳	منابع و مراجع
۵۱	پیوست
۵۲	واژه‌نامه‌ی فارسی به انگلیسی
۵۴	واژه‌نامه‌ی انگلیسی به فارسی

## فهرست شکل‌ها

صفحه

شکل

۳۹	۱-۶ دانشگاه امیرکبیر
۳۹	۲-۶ پلی تکنیک

## فهرست جدول‌ها

صفحه	جدول
۴۰	۱-۶ مقایسه الگوریتم‌های بهینه‌ساز
۴۱	۲-۶ پارامترهای الگوریتم
۴۱	۳-۶ پارامترهای الگوریتم
۴۱	۴-۶ پارامترهای الگوریتم

## فهرست نمادها

نماد	مفهوم
$\mathbb{R}^n$	فضای اقلیدسی با بعد $n$
$\mathbb{S}^n$	کره $n$ بعدی
$M^m$	خمینه $m$ -بعدی $M$
$\mathfrak{X}(M)$	جبر میدان‌های برداری هموار روی $M$
$\mathfrak{X}^1(M)$	مجموعه میدان‌های برداری هموار یک‌ه روی $(M, g)$
$\Omega^p(M)$	مجموعه $p$ -فرمی‌های روی خمینه $M$
$Q$	اپراتور ریچی
$\mathcal{R}$	تانسور انحنای ریمان
$ric$	تانسور ریچی
$L$	مشتق لی
$\Phi$	۲-فرم اساسی خمینه تماسی
$\nabla$	التصاق لوی-چویتای
$\Delta$	لاپلاسین ناهموار
$\nabla^*$	عملگر خودالحاق صوری القا شده از التصاق لوی-چویتای
$g_s$	متر ساساکی
$\nabla$	التصاق لوی-چویتای وابسته به متر ساساکی
$\Delta$	عملگر لاپلاس-بلترامی روی $p$ -فرم‌ها

## فصل اول

راهنمای استفاده از الگوی لاتک دانشگاه

صنعتی امیرکبیر (پلی تکنیک تهران)

## ۱-۱ مقدمه

حروف چینی پروژه کارشناسی، پایان نامه یا رساله یکی از موارد پرکاربرد استفاده از زی پرشین است. از طرفی، یک پروژه، پایان نامه یا رساله، احتیاج به تنظیمات زیادی از نظر صفحه آرایی دارد که ممکن است برای یک کاربر مبتدی، مشکل باشد. به همین خاطر، برای راحتی کار کاربر، یک کلاس با نام AUTthesis برای حروف چینی پروژه ها، پایان نامه ها و رساله های دانشگاه صنعتی امیرکبیر با استفاده از نرم افزار زی پرشین، آماده شده است. این فایل به گونه ای طراحی شده است که کلیه خواسته های مورد نیاز مدیریت تحصیلات تکمیلی دانشگاه صنعتی امیرکبیر را برآورده می کند و نیز، حروف چینی بسیاری از قسمت های آن، به طور خودکار انجام می شود.

کلیه فایل های لازم برای حروف چینی با کلاس گفته شده، داخل پوشه ای به نام AUTthesis قرار داده شده است. توجه داشته باشید که برای استفاده از این کلاس باید فونت های PGaramond, B Nazanin و IranNastaliq روی سیستم شما نصب شده باشد.

## ۲-۱ این همه فایل؟!

از آنجایی که یک پایان نامه یا رساله، یک نوشته بلند محسوب می شود، لذا اگر همه تنظیمات و مطالب پایان نامه را داخل یک فایل قرار بدهیم، باعث شلوغی و سردرگمی می شود. به همین خاطر، قسمت های مختلف پایان نامه یا رساله داخل فایل های جداگانه قرار گرفته است. مثلاً تنظیمات پایه ای کلاس، داخل فایل AUTthesis.cls، تنظیمات قابل تغییر توسط کاربر، داخل commands.tex، قسمت مشخصات فارسی پایان نامه، داخل fa\_title.tex، مطالب فصل اول، داخل chapter1 و ... قرار داده شده است. نکته مهمی که در اینجا وجود دارد این است که از بین این فایل ها، فقط فایل AUTthesis.tex قابل اجرا است. یعنی بعد از تغییر فایل های دیگر، برای دیدن نتیجه تغییرات، باید این فایل را اجرا کرد. بقیه فایل ها به این فایل، کمک می کنند تا بتوانیم خروجی کار را ببینیم. اگر به فایل AUTthesis.tex دقت کنید، متوجه می شوید که قسمت های مختلف پایان نامه، توسط دستورهایی مانند input و include به فایل اصلی، یعنی AUTthesis.tex معرفی شده اند. بنابراین، فایلی که همیشه با آن سروکار داریم، فایل AUTthesis.tex است. در این فایل، فرض شده است که پایان نامه یا رساله شما، از ۵ فصل و یک پیوست، تشکیل شده است. با این حال، اگر پایان نامه یا رساله شما، بیشتر از ۵ فصل و یک پیوست است، باید خودتان فصل های بیشتر را به این فایل، اضافه کنید. این کار، بسیار ساده است. فرض کنید بخواهید

یک فصل دیگر هم به پایان‌نامه، اضافه کنید. برای این کار، کافی است یک فایل با نام chapter6 و با پسوند .tex بسازید و آن را داخل پوشه AUTthesis قرار دهید و سپس این فایل را با دستور `\include{chapter6}` داخل فایل AUTthesis.tex و بعد از دستور `\include{chapter6}` قرار دهید.

## ۳-۱ از کجا شروع کنم؟

قبل از هر چیز، بدیهی است که باید یک توزیع تک مناسب مانند TeX Live و یک ویرایش‌گر تک مانند Texmaker را روی سیستم خود نصب کنید. نسخه بهینه شده Texmaker را می‌توانید از سایت [پارسی لاتک<sup>۱</sup>](http://www.parsilatex.com) و TeX Live را هم می‌توانید از [سایت رسمی آن<sup>۲</sup>](http://www.tug.org/texlive) دانلود کنید.

در مرحله بعد، سعی کنید که یک پشتیبان از پوشه AUTthesis بگیرید و آن را در یک جایی از هارددیسک سیستم خود ذخیره کنید تا در صورت خراب کردن فایل‌هایی که در حال حاضر، با آن‌ها کار می‌کنید، همه چیز را از دست ندهید.

حال اگر نوشتن پایان‌نامه اولین تجربه شما از کار با لاتک است، توصیه می‌شود که یک‌بار به طور سری، کتاب «مقدمه‌ای نه چندان کوتاه بر  $\text{\LaTeX}$ <sup>۳</sup>» ترجمه دکتر مهدی امیدعلی، عضو هیات علمی دانشگاه شاهد را مطالعه کنید. این کتاب، کتاب بسیار کاملی است که خیلی از نیازهای شما در ارتباط با حروف‌چینی را برطرف می‌کند.

بعد از موارد گفته شده، فایل AUTthesis.tex و fa\_title را باز کنید و مشخصات پایان‌نامه خود مثل نام، نام خانوادگی، عنوان پایان‌نامه و ... را جایگزین مشخصات موجود در فایل fa\_title کنید. دقت داشته باشید که نیازی نیست نگران چینش این مشخصات در فایل پی‌دی‌اف خروجی باشید. فایل AUTthesis.cls همه این کارها را به طور خودکار برای شما انجام می‌دهد. در ضمن، موقع تغییر دادن دستورهای داخل فایل fa\_title کاملاً دقت کنید. این دستورها، خیلی حساس هستند و ممکن است با یک تغییر کوچک، موقع اجرا، خطا بگیرید. برای دیدن خروجی کار، فایل fa\_title را Save، (نه Save As) کنید و بعد به فایل AUTthesis.tex برگشته و آن را اجرا کنید. حال اگر می‌خواهید مشخصات انگلیسی پایان‌نامه را هم عوض کنید، فایل en\_title را باز کنید و مشخصات

<sup>۱</sup><http://www.parsilatex.com>

<sup>۲</sup><http://www.tug.org/texlive>

<sup>۳</sup><http://www.tug.ctan.org/tex-archive/info/lshort/persian/lshort.pdf>



داخل آن را تغییر دهید.<sup>۴</sup> در اینجا هم برای دیدن خروجی، باید این فایل را Save کرده و بعد به فایل AUTthesis.tex برگشته و آن را اجرا کرد.

برای راحتی بیشتر، فایل AUTthesis.cls طوری طراحی شده است که کافی است فقط یکبار مشخصات پایان‌نامه را وارد کنید. هر جای دیگر که لازم به درج این مشخصات باشد، این مشخصات به طور خودکار درج می‌شود. با این حال، اگر مایل بودید، می‌توانید تنظیمات موجود را تغییر دهید. توجه داشته باشید که اگر کاربر مبتدی هستید و یا با ساختار فایل‌های cls آشنایی ندارید، به هیچ وجه به این فایل، یعنی فایل AUTthesis.cls دست نزنید.

نکته دیگری که باید به آن توجه کنید این است که در فایل AUTthesis.cls، سه گزینه به نام‌های msc, bsc و phd برای تایپ پروژه، پایان‌نامه و رساله، طراحی شده است. بنابراین اگر قصد تایپ پروژه کارشناسی، پایان‌نامه یا رساله را دارید، در فایل AUTthesis.tex باید به ترتیب از گزینه‌های msc, bsc و phd استفاده کنید. با انتخاب هر کدام از این گزینه‌ها، تنظیمات مربوط به آنها به طور خودکار، اعمال می‌شود.

## ۴-۱ مطالب پایان‌نامه را چگونه بنویسم؟

### ۱-۴-۱ نوشتن فصل‌ها

همان‌طور که در بخش ۱-۲ گفته شد، برای جلوگیری از شلوغی و سردرگمی کاربر در هنگام حروف‌چینی، قسمت‌های مختلف پایان‌نامه از جمله فصل‌ها، در فایل‌های جداگانه‌ای قرار داده شده‌اند. بنابراین، اگر می‌خواهید مثلاً مطالب فصل ۱ را تایپ کنید، باید فایل‌های AUTthesis.tex و chapter1 را باز کنید و محتویات داخل فایل chapter1 را پاک کرده و مطالب خود را تایپ کنید. توجه کنید که همان‌طور که قبلاً هم گفته شد، تنها فایل قابل اجرا، فایل AUTthesis.tex است. لذا برای دیدن حاصل (خروجی) فایل خود، باید فایل chapter1 را Save کرده و سپس فایل AUTthesis.tex را اجرا کنید. یک نکته بدیهی که در اینجا وجود دارد، این است که لازم نیست که فصل‌های پایان‌نامه را به ترتیب تایپ کنید. می‌توانید ابتدا مطالب فصل ۳ را تایپ کنید و سپس مطالب فصل ۱ را تایپ کنید.

نکته بسیار مهمی که در اینجا باید گفته شود این است که سیستم  $\text{\TeX}$ ، محتویات یک فایل تک

---

<sup>۴</sup> برای نوشتن پروژه کارشناسی، نیازی به وارد کردن مشخصات انگلیسی پروژه نیست. بنابراین، این مشخصات، به طور خودکار، نادیده گرفته می‌شود.

را به ترتیب پردازش می‌کند. به عنوان مثال، اگر فایلی، دارای ۴ خط دستور باشد، ابتدا خط ۱، بعد خط ۲، بعد خط ۳ و در آخر، خط ۴ پردازش می‌شود. بنابراین، اگر مثلاً مشغول تایپ مطالب فصل ۳ هستید، بهتر است که دو دستور `\include{chapter1}` و `\include{chapter2}` را در فایل `AUTthesis.tex` غیرفعال<sup>۵</sup> کنید. زیرا در غیر این صورت، ابتدا مطالب فصل ۱ و ۲ پردازش شده (که به درد ما نمی‌خورد؛ چون ما می‌خواهیم خروجی فصل ۳ را ببینیم) و سپس مطالب فصل ۳ پردازش می‌شود و این کار باعث طولانی شدن زمان اجرا می‌شود. زیرا هر چقدر حجم فایل اجرا شده، بیشتر باشد، زمان بیشتری هم برای اجرای آن، صرف می‌شود.

## ۲-۴-۱ مراجع

برای وارد کردن مراجع به فصل ۲ مراجعه کنید.

## ۳-۴-۱ واژه‌نامه فارسی به انگلیسی و برعکس

برای وارد کردن واژه‌نامه فارسی به انگلیسی و برعکس، بهتر است مانند روش بکار رفته در فایل‌های `dicen2fa` و `dicfa2en` عمل کنید.

## ۵-۱ اگر سوالی داشتیم، از کی بپرسیم؟

برای پرسیدن سوال‌های خود در مورد حروف چینی با زی‌پرشین، می‌توانید به **تالار گفتگوی پارسی لاتک**<sup>۶</sup> مراجعه کنید. شما هم می‌توانید روزی به سوال‌های دیگران در این تالار، جواب بدهید.

---

<sup>۵</sup> برای غیرفعال کردن یک دستور، کافی است پشت آن، یک علامت `%` بگذارید.

<sup>۶</sup><http://www.forum.parsilatex.com>

## فصل دوم

### مرور کارهای پیشین

در این فصل به معرفی یادگیری پیوسته، یادگیری پیوسته در دو حوزه ی تصویر و ویدیو و مدل های بنیایی-زبان در این حوزه ها می پردازیم.

## ۱-۲ یادگیری پیوسته

یادگیری پیوسته به توانایی یک عامل هوشمند یا سامانه برای کسب، به روزرسانی، جمع آوری و بهره برداری از دانش در طول عمر آن اشاره دارد. این، شامل یادگیری یک دنباله از مطالب یا وظایف یکی پس از دیگری و سازگاری با اطلاعات جدید بدون فراموشی دانش قبلاً آموخته شده است. به علت تدریجی اضافه شدن مطالب، یادگیری مداوم به عنوان یادگیری افزایشی<sup>۱</sup>، یادگیری مستمر<sup>۲</sup> و یادگیری مادام العمر<sup>۳</sup> نیز شناخته می شود. هدف یادگیری پیوسته حفظ تعادل بین یادگیری اطلاعات جدید و حفظ دانش قبلاً کسب شده، با غلبه بر چالش فراموشی فاجعه بار است. در ادامه به معرفی چالش ها، سناریوها، رویکردها و کاربردهای این حوزه می پردازیم.

### ۱-۱-۲ چالش ها

مسئله ی اصلی که در این نوع یادگیری به وجود می آید، چالشی با نام فراموشی فاجعه بار است. علت این امر این است که وقتی داده های جدید برای یادگیری اضافه می شوند، سامانه مجبور می شود اطلاعات قبل را تا حدی به فراموشی بسپرد و شرایط جدید را در نظر بگیرد. برای همین به نوعی یک تعادل بین حالتی که حافظه ثابت است و حالتی که یادگیری انعطاف پذیر است، نیاز میباشد. در واقع هدف بر این است که یادگیری مستمر تعمیم پذیری خوبی برای تطبیق در شرایط مختلف با داده های جدید و با توزیع های جدید داشته باشد، همچنین کارایی منابع را نیز تضمین کند [۱، ۲].

### ۲-۱-۲ سناریوها

بر اساس نوع داده ها و وظیفه هایی که باید انجام شود، سناریوهای مختلفی برای یادگیری پیوسته ارائه شده است [۱]:

<sup>1</sup>Incremental learning

<sup>2</sup>Continuous learning

<sup>3</sup>Lifelong learning

## یادگیری نمونه‌ای افزایشی

در یادگیری نمونه‌ای-افزاینده<sup>۴</sup> نمونه‌های داده‌ی جدید به طور پیوسته به مدل معرفی می‌شوند که هر یک نقطه داده جدید را نشان می‌دهد. این مدل باید با توزیع داده‌های در حال تکامل سازگار شود و در عین حال نمونه‌های جدید را نیز در نظر بگیرد.

## یادگیری افزایشی دامنه

در یادگیری افزایشی دامنه<sup>۵</sup>، دامنه تغییر می‌کند و به معنای افزایش یا تغییر در توزیع داده‌ها، ویژگی‌ها یا محیط‌ها در مسئله‌ای که مدل در حال یادگیری آن است، می‌باشد. این تغییر می‌تواند به صورت افزودن داده‌های جدید از دامنه‌های جدید، تغییر در ویژگی‌ها، یا حتی تغییر مفهوم برخی اجزاء از داده‌ها (مثلاً تغییر تعبیر برچسب‌ها) اتفاق بیافتد. چالش این است که اطمینان حاصل شود که مدل می‌تواند با این حوزه‌های جدید سازگار شود، بدون اینکه عملکرد آن در حوزه‌هایی که قبلاً دیده شده‌اند، کاهش یابد [۲، ۳، ۴].

## یادگیری افزایشی وظیفه

در یادگیری افزایشی وظیفه<sup>۶</sup> وظایف جدید در طول زمان معرفی می‌شوند و مدل باید یاد بگیرد که در هر وظیفه‌ی جدید به خوبی عمل کند و در عین حال عملکرد خود را در وظایفی که قبلاً آموخته‌است حفظ کند [۳].

## یادگیری افزایشی دسته‌ای

دسته‌های جدید به تدریج به داده‌های آموزشی مدل اضافه می‌شوند. مدل باید بیاموزد که دسته‌های جدید را تشخیص دهد و بین آنها تمایز قائل شود بدون اینکه دانش خود را در مورد دسته‌های آموخته شده قبلی فراموش کند [۲، ۳، ۵، ۶].

## ۲-۱-۳ رویکردها

وَنگ و همکاران [۱]، یادگیری پیوسته را به صورت جامع به دسته‌های زیر تقسیم کرده‌اند:

<sup>4</sup>Instance incremental learning

<sup>5</sup>Domain incremental learning

<sup>6</sup>Task incremental learning

### رویکرد مبتنی بر تنظیم

این رویکرد بر اساس این است که از تکنیک های مبتنی بر تنظیم<sup>۷</sup> برای ایجاد تعادل در وظایف قدیم و جدید استفاده کند. در حالت کلی نیز تکنیک های منظم سازی با ایجاد تغییراتی در محاسبات باعث جلوگیری از بیش برآزش مدل می شوند و در این جا نیز هدف این است که بین تاثیر مدل های قبلی و جدید تعادل ایجاد کند. برای رسیدن به این هدف نیز نیاز است که یک کپی از مدل های قبلی داشته باشد تا باعث فراموشی نشود. بر اساس اینکه چه نوع روشی استفاده شود، منظم سازی به دو دسته تقسیم می شود: منظم سازی وزن ( که وزن هایی که اهمیت بیشتری در نتیجه دارند را بیشتر نگه می دارد و وزن های بی اهمیت به مدل جدید با ضریب خیلی کم یا صفر انتقال می یابند). و منظم سازی تابع ( مدل جدید که با نام شاگرد از آن یاد می کند، تلاش می کند از اطلاعات خروجی مدل قبلی که با نام معلم از آن یاد می شود، برای ایجاد خروجی وظیفه ی خودش راهنمایی بگیرد) [۷، ۸].

### رویکرد مبتنی بر بازپخش

هدف این رویکرد در واقع این است که بخشی از داده های قبلی را در حافظه ذخیره کرده و آن ها را در وظیفه ی جدید با داده های جدید آموزش دهد و به این صورت هم مدل جدید آموزش داده می شود و هم اطلاعات قبلی فراموش نمی شوند. این رویکرد نیز به سه دسته ی بازپخش تجربه<sup>۸</sup> (انتقال بخشی از داده های قبلی به حافظه)، بازپخش مولد<sup>۹</sup> (ایجاد یک مدل مولد اضافی برای بازپخش داده های تولید شده)، بازپخش ویژگی<sup>۱۰</sup> (انتقال ویژگی های مهم داده های قبلی به وظیفه ی جدید و جلوگیری از فراموشی فاجعه بار). به عنوان مثال شین و همکاران [۹]، مدلی به نام مدل مولد عمیق ارائه داده اند که چارچوبی است از یک معماری مدل دوگانه تعاونی، با یک مدل مولد عمیق (مولد) و یک مدل حل وظیفه (حل کننده) تشکیل شده است. مولد، مسئول تولید ورودی های جعلی است که شبیه داده های گذشته است، در حالی که حل کننده برای حل وظیفه فعلی آموزش دیده است. با درهم آمیختن داده های آموزشی برای وظایف قبلی با داده های مربوط به وظیفه جدید، مدل می تواند بدون فراموش کردن دانش وظایف قدیمی، وظایف جدید را یاد بگیرد. این رویکرد از ماهیت مولد هیپوکامپ، یک سامانه حافظه کوتاه مدت در مغز پستانداران الهام گرفته شده است.

<sup>7</sup>Regularization based

<sup>8</sup>Experience replay

<sup>9</sup>Generative replay

<sup>10</sup>Feature replay

### رویکرد مبتنی بر بهینه سازی

رویکرد مبتنی بر بهینه سازی<sup>۱۱</sup> به جای تغییر در تابع خطا، به دستکاری برنامه های بهینه سازی می پردازد مثلاً به روزرسانی پارامترها به گونه ای صورت گیرد که با جهت هایی مانند فضای ورودی قبلی عمود یا هم تراز باشند تا اطلاعات وظایف قبلی نیز حفظ شود.

### رویکرد مبتنی بر معماری

رویکرد مبتنی بر معماری<sup>۱۲</sup> به ساخت پارامترهای خاص هر وظیفه برای جلوگیری از دخالت وظیفه ها در یکدیگر می پردازد و در واقع هدف بر کاهش محدودیت استفاده از پارامترهای مشترک وظیفه هاست زیرا با وجود این محدودیت ها، اطلاعات کمتری از وظیفه قبل به بعد منتقل می شود و فراموشی فاجعه بار با احتمال بیشتری رخ می دهد. در این رویکرد نیز تکنیک های مختلفی ارائه شده است مانند تخصیص پارامتر، تجزیه مدل و شبکه پیمانه ای.

### رویکرد حاصل از ترکیب رویکردها و سناریوها

یکی از راه حل ها این است که از ترکیب رویکردهای ذکر شده با یکدیگر یا با دیگر رویکردهای شبکه عصبی استفاده شود. مثلاً استفاده از ترکیب رویکردهای مبتنی بر منظم سازی و مبتنی بر بازپخش.

## ۴-۱-۲ کاربردها

یادگیری پیوسته در حوزه بینایی ماشین کاربردهای گسترده ای دارد که از مهم ترین آن ها می توان به تشخیص چهره، دسته بندی تصویر، تشخیص اعمال در ویدیو، بخش بندی معنایی و تلفیق زبان و بینایی اشاره کرد. این قابلیت ها امکان آموزش پیوسته مدل ها را بدون فراموشی اطلاعات قبلی فراهم می سازند و باعث می شوند سامانه های هوشمند در مواجه شدن با داده های جدید، ضمن حفظ دانش گذشته، عملکرد خود را ارتقا دهند.

---

<sup>11</sup>Optimization based approach

<sup>12</sup>Architecture based

## ۲-۲ یادگیری پیوسته در بینایی کامپیوتر

یادگیری پیوسته در زمینه‌های مختلفی مورد استفاده قرار گرفته است. یکی از این زمینه‌ها بینایی کامپیوتر است که دو حوزه کاربرد به نام دسته‌بندی تصویر و تشخیص عمل را شامل شده و در ادامه به بررسی مقالات مطرح درباره آن‌ها می‌پردازیم.

### ۱-۲-۲ دسته‌بندی تصویر

مای و همکاران [۲]، مقایسه‌ای بین رویکردهای مطرح ارائه شده برای دسته‌بندی تصویر انجام داده‌اند که در ادامه به معرفی مختصر رویکردها و مقایسه آن‌ها می‌پردازیم.

#### روش تثبیت وزن کشسان

جیمز و همکاران [۷] مقاله‌ای در رابطه با روش جدیدی در زمینه‌ی منظم‌سازی ارائه داده‌اند. تثبیت وزن کشسان<sup>۱۳</sup> یک الگوریتم است که به شبکه‌های عصبی عمیق اجازه می‌دهد تا مجموعه‌ای از وظایف پیچیده را بدون فراموش کردن فاجعه‌آمیز یاد بگیرند. این کار را با کاهش انتخابی انعطاف‌پذیری وزن انجام می‌دهد. این روش با محدود کردن هر وزن با یک جریمه درجه دوم کار می‌کند که آن را به مقداری متناسب با اهمیت آن برای عملکرد در وظایفی که قبلاً یاد گرفته شده، به سمت مقادیر قدیمی خود می‌کشاند. در واقع از ماتریسی به نام ماتریس اطلاعات فشر برای محاسبه‌ی اهمیت وزن‌ها در مدل قبلی استفاده می‌کند.

#### روش یادگیری بدون فراموشی

لی و همکاران [۸]، الگوریتم یادگیری بدون فراموشی<sup>۱۴</sup> را ارائه داده‌اند. این الگوریتم برای یادگیری بدون فراموشی یک الگوریتم یادگیری پیوسته است که هدف آن یادگیری وظایف جدید بدون فراموشی دانش وظایف قبلی می‌باشد. با آموزش یک مدل اولیه بر روی یک مجموعه وظایف اولیه کار می‌کند. این مدل اولیه به عنوان مدل استاد نامیده می‌شود. هنگامی که با یک وظیفه جدید مواجه می‌شود، یک مدل جدید با نام مدل شاگرد، با استفاده از داده‌های وظیفه جدید، آموزش می‌بیند. مدل شاگرد همچنین با استفاده از داده‌های وظیفه‌های قدیمی، آموزش می‌بیند تا از مدل استاد پیروی کند. این به صورت

<sup>13</sup>Elastic weight consolidation (EWC)

<sup>14</sup>Learning without forgetting (LWF)



محاسبه خروجی مدل استاد برای داده‌های جدید و سپس استفاده از آن خروجی به عنوان هدف برای آموزش مدل شاگرد انجام می‌شود. این روند به مدل شاگرد کمک می‌کند تا دانش وظایف قدیمی را حفظ کند، در حالی که در عین حال برای وظیفه جدید نیز بهینه می‌شود. از این روند به عنوان تقطیر دانش<sup>۱۵</sup> نیز یاد می‌کنند. عملکرد آن می‌تواند در صورتی که وظیفه جدید بسیار متفاوت از وظایف قدیمی باشد، کاهش یابد.

### روش میانگین حافظه رخدادی گرادیان

چودری و همکاران [۱۰]، الگوریتمی به نام الگوریتم میانگین حافظه رخدادی گرادیان<sup>۱۶</sup> ارائه کردند که با استفاده از یک حافظه ذخیره شده برای ذخیره اطلاعات مربوط به وظایف قبلی، از فراموشی فاجعه بار جلوگیری می‌کند. در هر مرحله از یادگیری، مدل از حافظه ذخیره شده برای تولید یک زیرمجموعه تصادفی از تجارب استفاده می‌کند. این تجربیات سپس برای محاسبه گام‌های گرادیان استفاده می‌شوند. گام‌های گرادیان فعلی سپس با گام‌های گرادیان زیرمجموعه تصادفی از حافظه ذخیره شده میانگین گیری می‌شوند. این میانگین گیری به مدل کمک می‌کند تا یک دیدگاه کلی تر از وظیفه فعلی به دست آورد. این دیدگاه کلی تر به مدل کمک می‌کند تا دانش وظایف قبلی خود را حفظ کند، حتی زمانی که در حال یادگیری یک وظیفه جدید است.

### روش طبقه بندی افزایشی و یادگیری بازنمایی

ربوفی و همکاران [۱۱]، روش طبقه بندی افزایشی و یادگیری بازنمایی<sup>۱۷</sup> در یادگیری پیوسته ارائه دادند که مدل با استفاده از نمونه‌هایی از همه دسته‌ها، از جمله دسته‌های جدید و قدیمی، آموزش داده می‌شود. تابع ضرر شامل یک ضرر طبقه بندی برای تشویق مدل به پیش بینی برچسب‌های صحیح برای دسته‌های جدید و یک ضرر تقطیر دانش برای تشویق مدل به بازتولید خروجی‌های مدل قبلی برای دسته‌های قدیمی است. همچنین یک روش به روزرسانی حافظه را پیشنهاد می‌دهد که بر اساس فاصله در فضای ویژگی‌های نهفته است. این روش برای انتخاب زیرمجموعه‌ای از نمونه‌ها از هر دسته استفاده می‌شود که میانگین ویژگی‌های نهفته آنها به میانگین همه نمونه‌ها در این دسته نزدیک‌ترین است.

<sup>15</sup>Knowledge distillation (KD)

<sup>16</sup>Averaged Gradient Episodic Memory (A-GEM)

<sup>17</sup>Incremental Classifier and Representation Learning (iCaRL)

## روش حداکثر تداخل ارزیابی

الجانندی و همکاران [۱۲]، روشی به نام حداکثر تداخل ارزیابی<sup>۱۸</sup> را ارائه داده اند که یک روش مبتنی بر بازپخش است که اخیراً با هدف بهبود راهبرد بازیابی حافظه پیشنهاد شده است. این روش، نمونه‌های بازپخش را با توجه به افزایش ضرر با داشتن به به‌روزرسانی پارامتر تخمین زده شده بر اساس دسته‌های کوچک ورودی انتخاب می‌کند. نمونه‌های حافظه را که بیشترین تداخل (افزایش ضرر) را با به‌روزرسانی پارامتر با دسته ورودی جدید دارند، انتخاب می‌کند. همچنین نمونه‌گیری مخزن را در به‌روزرسانی حافظه اعمال می‌کند و نمونه‌های حافظه انتخابی را با نمونه‌های جدید در به‌روزرسانی مدل دوباره پخش می‌کند.

## ۲-۲-۲ تشخیص عمل

تشخیص عمل یکی از مباحث پیشرفته امروزی است که به علت اضافه شدن پارامتر زمان، پیچیدگی‌های بیشتری نسبت به تصویر پیدا کرده است. همچنین به علت حجم بودن داده‌ها در این مسائل، یادگیری مداوم ضرورت پیدا می‌کند. زیرا در طول زمان مثلاً دسته‌ها و داده‌های بیشتری به مدل اضافه می‌شوند و مدل نمی‌تواند دوباره از ابتدا این حجم داده را آموزش دهد. پس چند سال اخیر، مطالعه‌هایی نیز در این زمینه شده و روش‌های جدید با مجموعه داده‌های مختلف ارائه شده است. که به چندین مقاله در ادامه می‌پردازیم.

## روش مین هاس

مین هاس و همکاران [۱۳]، یک روش یادگیری پیوسته برای تشخیص اعمال انسان ارائه داده‌اند. این روش بر پایه یک چارچوب است که دو تکنیک، یعنی تقریب شکل و یادگیری تحلیلی، را با هم ترکیب می‌کند. تقریب شکل برای ثبت شکل بازیگر در ویدیو استفاده می‌شود. به این صورت که با تغییراتی که به شکل می‌دهیم، از تصاویر شدت نور بهره‌برداری می‌کنیم تا ویژگی‌های مربوط به جهت گرادیان‌ها را استخراج کنیم. هنگامی که حرکت در ویدیو پیش می‌رود، شکل با تغییر و تنظیم چندین قطعه کوچک داخل یک پنجره‌ی پیگیری، به‌طور دقیق به تغییرات مرزها پیگیری می‌کند. به منظور یادگیری دینامیک‌های غیرخطی حرکت‌ها، از یادگیری تحلیلی استفاده می‌شود. این فرآیند یادگیری به شکل بازگشتی انجام می‌شود و از طریق آن آموزش به نمایش خطی ساده تبدیل می‌شود. این روش دو مزیت دارد: کمینه کردن خطاها و کاهش قابل توجه زمان محاسباتی، و از بین بردن محدودیت‌های آموزش

<sup>18</sup>Maximally Interfered Retrieval (MIR)

به صورت دسته‌ای برای تشخیص اعمال. این روش یادگیری پیوسته اجازه می‌دهد که مدل به تدریج با ورودی داده‌های جدید به‌روز شود. این روش مقابل یادگیری دسته‌ای است که در آن برای آموزش دسته‌بند تمام مجموعه داده آموزشی استفاده می‌شود.

## روش الزنت

لی و همکاران [۱۴] روشی به نام الزنت را ارائه کردند که با انتخاب و به‌روزرسانی پویاترین بلوک‌های یادگیری از فراموشی فاجعه‌بار در شناسایی عمل جلوگیری می‌کند. هنگام یادگیری اعمال جدید، الزنت به دنبال بلوک‌های یادگیری می‌گردد که بیشترین ارتباط را با عمل فعلی دارند و پارامترهای آن‌ها را به‌روزرسانی می‌کند، در حالی که پارامترهای بلوک‌های غیرانتخابی را حفظ می‌کند. این راهبرد به‌روزرسانی انتخابی به حفظ دانش حرکات قبلاً یادگیری شده کمک می‌کند و مشکل فراموشی را کاهش می‌دهد. با به‌روزرسانی فقط بلوک‌های مرتبط، از وارد کردن نویز و اختلالات غیرمرتبط به دانش قبلی جلوگیری می‌کند، که منجر به عملکرد بهتر در یادگیری اعمال جدید می‌شود.

## روش تعبیه همسایه تی تصادفی موقت تحت نظارت

چنگ و همکاران [۱۵] یک روش برای تشخیص حرکات انسان با استفاده از تعبیه‌سازی همسایگی تصادفی زمانی نظارت شده و یادگیری پیوسته ارائه کرده‌اند. الگوریتم برای یادگیری ارتباط بین قاب‌های عمل به کار می‌رود و اطلاعات دسته و زمانی را تلفیق می‌کند. یادگیری پیوسته برای تعبیه‌سازی کم‌بعدی داده‌های جدید با استفاده از رویکردهایی نظیر تعبیه خطی محلی و پیش بینی حفظ محلی استفاده می‌شود. همچنین سه روش برای یادگیری پیوسته در زمینه تشخیص حرکات انسان توصیف می‌کند.

## روش پاریزی

پاریزی و همکاران [۱۶] رویکردی را ارائه کرده‌اند که باعث جلوگیری از فراموشی دانش با شبکه‌ی سلسله مراتبی خودسازمانده می‌شود. در این شبکه‌ی سلسله مراتبی هر لایه به صورت شبکه رشد هنگام نیاز است به این صورت که نورون‌های جدید را تخصیص می‌دهد یا نورون‌های موجود را بر اساس اختلاف بین توزیع ورودی و وزن‌های نورون‌های نمونه‌ای به‌روز می‌کند.

## ۳-۲ مدل‌های بینایی-زبان

مدل‌های زبانی بزرگ<sup>۱۹</sup>، شبکه‌های عصبی ترنسفورمر مقیاس‌پذیری هستند که با آموزش بر حجم عظیمی از داده‌های متنی، قادرند زبان طبیعی را تولید، درک و تحلیل کنند. این مدل‌ها به دلیل ظرفیت بالای خود در یادگیری، نقش اساسی در پیشرفت‌های اخیر پردازش زبان طبیعی<sup>۲۰</sup> ایفا کرده‌اند. به دنبال این پیشرفت‌ها پژوهش‌های زیادی انجام شده است که از این مدل‌ها در کاربردهای بینایی ماشین نیز استفاده کرده‌اند [۱۷]. مدل‌های بینایی-زبان<sup>۲۱</sup> دسته‌ای از مدل‌های هوش مصنوعی هستند که به طور هم‌زمان قادر به تحلیل و درک داده‌های بصری (تصویر یا ویدیو) و زبانی (متن) می‌باشند. این مدل‌ها با استفاده از حجم انبوهی از داده‌های تصویر-متن که به صورت گسترده در وب موجود است، آموزش می‌بینند. ایده اصلی پشت این مدل‌ها، یادگیری هم‌بستگی میان نمایش‌های تصویری و متنی در یک فضای مشترک نهفته<sup>۲۲</sup> است. به عنوان نمونه، مدل CLIP<sup>۲۳</sup> [۱۸] که توسط OpenAI ارائه شده است، با بهره‌گیری از صدها میلیون جفت تصویر و متن، توانسته است عملکرد قابل قبولی در وظایف مختلف بینایی و زبانی ارائه دهد. مدل‌های بینایی-زبان به دلایل متعددی مورد توجه پژوهشگران قرار گرفته‌اند که به برخی از آن‌ها در ادامه اشاره می‌کنیم [۱۷]:

- توانایی پیش‌بینی در حالت یادگیری بدون نمونه: این مدل‌ها قادرند وظایف جدید را بدون نیاز به بازآموزی<sup>۲۴</sup>، انجام دهند. به این حالت، یادگیری بدون نمونه<sup>۲۵</sup> گفته می‌شود.
- چندکاربردی بودن: یک مدل واحد می‌تواند در وظایف متنوعی همچون دسته‌بندی تصویر، تشخیص اشیاء، بازیابی تصویر بر اساس متن و تولید توضیح برای تصویر به کار گرفته شود.
- قابلیت مقیاس‌پذیری بالا: امکان آموزش بر روی میلیاردها جفت تصویر-متن و دستیابی به تعمیم‌پذیری قابل توجه در دامنه‌های گوناگون را دارد.

آموزش این مدل‌ها هزینه‌ی محاسباتی بالایی دارد اما طریقه‌ی استفاده از آن‌ها به صورتی است که این چالش تعدیل شود. استفاده از این مدل‌ها به سه مرحله‌ی اصلی تقسیم می‌شود: پیش‌آموزش، یادگیری

<sup>19</sup>Large Language Models (LLMs)

<sup>20</sup>Natural language processing

<sup>21</sup>Vision language models (VLMs)

<sup>22</sup>Embedding space

<sup>23</sup>Contrastive Language-Image Pre-training (CLIP)

<sup>24</sup>Retraining

<sup>25</sup>Zero-shot learning

انتقالی و تقطیر دانش که در ادامه بررسی می‌گردد.

## ۲-۳-۱ پیش آموزش مدل‌های بینایی-زبان

در مرحله‌ی پیش آموزش مدل‌های بینایی-زبان<sup>۲۶</sup> مدل با بهره‌گیری از حجم انبوهی از داده‌های تصویر-متن بدون برچسب، به‌گونه‌ای آموزش می‌بیند که توانایی درک هم‌زمان مفاهیم زبانی و تصویری را کسب کند. سه نوع هدف آموزشی عمده در این بخش عبارت‌اند از:

### اهداف تقابلی

در روش اهداف تقابلی<sup>۲۷</sup> مدل یاد می‌گیرد تا نمایش جفت‌های صحیح تصویر-متن را به یکدیگر نزدیک و جفت‌های نادرست را از هم دور کند. به عنوان مثال مدل CLIP که با هدف تقابلی و داده‌های وب‌مقیاس آموزش داده شد و توانست در بیش از ۳۰ وظیفه نمونه-صفر عملکرد موفق‌تری ارائه دهد.

### اهداف مولد

در اهداف مولد<sup>۲۸</sup> مدل به بازسازی بخش‌های حذف‌شده از تصویر یا متن می‌پردازد یا توصیف متنی برای تصویر تولید می‌کند. به عنوان نمونه، مدل FLAVA [۱۹] با بهره‌گیری هم‌زمان از ماسک‌کردن تصویر و زبان، دانش چندحالتی‌ای را در یک مدل واحد می‌آموزد.

### اهداف هم‌ترازی

اهداف هم‌ترازی<sup>۲۹</sup> بر هم‌خوانی معنایی میان تصویر و متن، به صورت کلی<sup>۳۰</sup> یا حتی به صورت محلی<sup>۳۱</sup> تمرکز دارند. مدل GLIP [۲۰] با هم‌ترازی زبان-ناحیه توانست به شناسایی اشیای واژگان-باز<sup>۳۲</sup> دست یابد.

<sup>26</sup>Vision-language model pre-training

<sup>27</sup>Contrastive objectives

<sup>28</sup>Generative objectives

<sup>29</sup>Alignment objectives

<sup>30</sup>image-text matching

<sup>31</sup>region-word matching

<sup>32</sup>Open-vocabulary

## ۲-۳-۲ یادگیری انتقالی مدل‌های بینایی-زبان

برای استفاده از مدل‌های بینایی-زبان در وظایف خاص مانند دسته‌بندی تصویر، تشخیص اشیا یا بازیابی تصویر، لازم است که مدل با روش‌هایی کم‌هزینه و تطبیقی انتقال یابد. مهم‌ترین روش‌ها عبارت‌اند از:

### تنظیم پرامپت

در روش تنظیم پرامپت<sup>۳۳</sup>، به جای تغییر ساختار داخلی مدل یا بازآموزی کامل آن، تلاش می‌شود تا ورودی‌های متنی (و در برخی موارد تصویری) به گونه‌ای هوشمندانه طراحی یا بهینه شوند که مدل بتواند عملکرد بهتری در وظیفه موردنظر ارائه دهد. در واقع، مدل اصلی ثابت می‌ماند و تنها شکل ورودی‌هایی که به آن داده می‌شود، به کمک الگوریتم‌هایی قابل یادگیری، تغییر می‌کند. این رویکرد به‌ویژه برای وظایف با یادگیری محدود بسیار کارآمد است؛ زیرا به مدل اجازه می‌دهد با استفاده از اطلاعات آموخته‌شده قبلی، خود را با وظیفه جدید تطبیق دهد بدون آنکه پارامترهایش خیلی تغییر کند. لیو و همکاران [۲۱]، تشخیص حرکت را به مسئله تطبیق ویدیو-متن تبدیل کرده‌اند تا از قدرت نمایش‌های زبانی بهره ببرند. پرامپت‌سازی نقش کلیدی در نزدیک‌سازی وظیفه هدف به ساختار داده‌های پیش‌تمرین شده دارد و موجب بهبود عملکرد مدل در شرایط یادگیری بدون نمونه می‌شود. در مدل CoOp [۲۲] به جای استفاده از پرامپت‌های متنی دستی مانند "a photo of a [class name]" پرامپت‌هایی از کلمات قابل یادگیری طراحی شدند. این کلمات به صورت بردارهایی آموزش‌پذیر به مدل داده می‌شوند و نقش آنها تقویت معنای دسته‌بندی برای مدل است. این روش، که در بخش بعدی بیشتر به آن پرداخته می‌شود، باعث شد دقت مدل CLIP در دسته‌بندی چنددسته به‌ویژه در شرایط یادگیری محدود به‌طور قابل توجهی بهبود یابد.

### وفق دهنده‌ی ویژگی

وفق دهنده‌ی ویژگی<sup>۳۴</sup> یکی از روش‌های مؤثر برای انتقال مدل‌های بینایی-زبان به وظایف جدید بدون نیاز به بازآموزی کامل شبکه است. در این رویکرد، به جای تغییر پارامترهای اصلی مدل، لایه‌هایی سبک و کم‌پارامتر به انتهای یا میانه‌ی شبکه اضافه می‌شود تا ویژگی‌های استخراج‌شده از تصویر یا متن را با نیازهای وظیفه‌ی خاص منطبق کند. این تطبیق‌دهنده‌ها می‌توانند به صورت افزونه‌هایی جدا از معماری اصلی عمل کنند، بنابراین هسته‌ی مدل بدون تغییر باقی می‌ماند. در روش CLIP-Adapter [۲۳]

<sup>33</sup>Prompt Tuning

<sup>34</sup>Feature adapter

مجموعه‌ای از لایه‌های سبک‌وزن به مدل CLIP افزوده شد تا ویژگی‌های استخراج‌شده از تصویر و متن پیش از تصمیم‌گیری نهایی پردازش و تطبیق یابند. این کار برای وظایف یادگیری کم‌نمونه نیز مؤثر بود، زیرا بدون نیاز به تغییر در مدل پایه، عملکرد بسیار مناسبی حاصل شد. این روش انتقالی به دلیل کم‌هزینه بودن و عدم نیاز به تنظیم مجدد کل مدل، برای بسیاری از کاربردهای عملی مناسب است.

## سایر روش‌ها

در کنار تنظیم پرامپت و تطبیق‌دهنده‌های ویژگی، برخی روش‌ها نیز با تغییر مستقیم در پارامترهای مدل، در بهبود عملکرد مدل برای وظایف خاص، نقش دارند. این روش‌ها معمولاً شامل تنظیم دقیق کامل یا تلفیق مدل‌های یادگرفته‌شده با مدل اولیه هستند. در روش Wise-FT [۲۴]، یک رویکرد ساده اما مؤثر ارائه شده است که در آن وزن‌های مدل پایه و مدل تنظیم دقیق شده به صورت میانگین‌گیری وزنی ترکیب می‌شوند. این تکنیک باعث می‌شود که مدل هم از تعمیم‌پذیری مدل اولیه بهره‌برد و هم بتواند دانش خاص وظیفه‌ی جدید را بیاموزد، بدون آنکه دچار بیش‌برازش شود. در توسعه‌ی این روش، ونگ و همکاران [۲۵]، روش Open-VCLIP را برای تطبیق مدل CLIP برای داده‌های ویدیویی ارائه دادند به صورتی که دانش مدل CLIP نیز حفظ شود. این مدل در فصل بعد به صورت مفصل‌تری توضیح داده خواهد شد.

## ۳-۳-۲ تقطیر دانش

در تقطیر دانش<sup>۳۵</sup> دانش مدل بینایی-زبان به یک مدل سبک‌تر منتقل می‌شود تا بتوان از آن در کاربردهای خاص و با منابع محدود استفاده کرد. دو کاربرد اصلی عبارت‌اند از:

### تقطیر دانش برای تشخیص شیء

در حوزه‌ی بینایی کامپیوتر، یکی از چالش‌های مهم، شناسایی اشیائی است که در داده‌های آموزش مدل پایه وجود نداشته‌اند. روش‌های متداول تشخیص شیء نیازمند برچسب‌گذاری دقیق و پرهزینه‌ی داده‌ها برای هر دسته هستند. در این میان، مدل‌های بینایی-زبان مانند CLIP که از داده‌های وب‌مقیاس و متن‌های توصیفی متنوع آموزش دیده‌اند، دارای دانش گسترده‌ای درباره‌ی مفاهیم بصری و زبانی هستند که می‌توان از آن‌ها برای توسعه‌ی مدل‌های تشخیص شیء استفاده کرد. در مدل VILD [۲۶] نمونه‌ای برجسته از این رویکرد است. این مدل با استفاده از تقطیر دانش از CLIP یک آشکارساز دو مرحله‌ای

<sup>35</sup>Knowledge distillation

توسعه داده است که می‌تواند اشیاء خارج از مجموعه‌ی برچسب‌گذاری‌شده‌ی اولیه را شناسایی کند؛ به این صورت که ویژگی‌های بصری استخراج‌شده از تصاویر با تعبیه‌های متنی مدل CLIP مقایسه می‌شوند تا به جای اتکا به دسته‌های از پیش تعریف‌شده، اشیاء جدید نیز قابل شناسایی باشند. این روش نوعی تشخیص شیء واژگان-باز را ممکن می‌سازد که در بسیاری از کاربردهای دنیای واقعی اهمیت بالایی دارد.

### تقطیر دانش برای بخش‌بندی معنایی

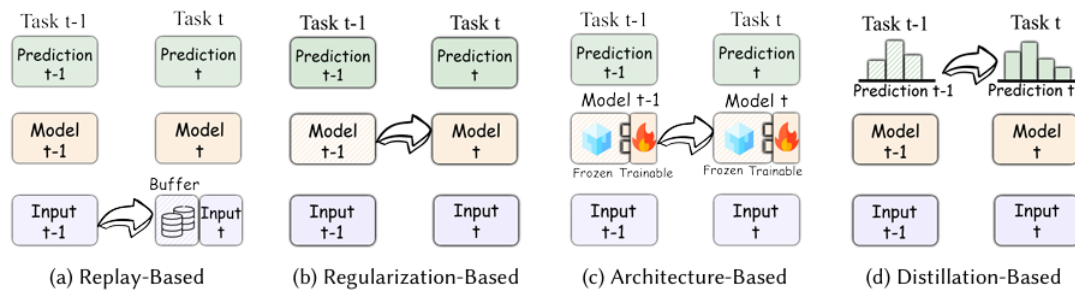
بخش‌بندی معنایی به معنای اختصاص یک برچسب معنایی به هر پیکسل از تصویر است و از وظایف کلیدی در درک صحنه محسوب می‌شود. پیاده‌سازی موفق این وظیفه معمولاً نیازمند مجموعه داده‌های پر حجم و برچسب‌خورده در سطح پیکسل است، که تولید آن‌ها بسیار پرهزینه و زمان‌بر است. با این حال، مدل‌های بینایی-زبان که از داده‌های ضعیف برچسب‌خورده یا بدون برچسب بهره می‌برند، می‌توانند دانش انتزاعی خود را به مدل‌های سبک‌تر انتقال دهند تا نیاز به برچسب‌گذاری کاهش یابد. در CLIPSeg [۲۷] از ویژگی‌های استخراج‌شده توسط CLIP برای هر تصویر استفاده می‌کند و با افزودن یک رمزگشای سبک<sup>۳۶</sup> امکان پیش‌بینی نقشه‌های بخش‌بندی معنایی را تنها بر اساس توصیف متنی (prompt) فراهم می‌کند؛ برای مثال، با دادن جمله‌ای مانند «گره در تصویر کجاست؟»، مدل قادر به تولید نقشه‌ای است که نواحی مربوط به گره را برجسته کند. نکته قابل توجه این است که این مدل به یادگیری بدون نمونه دست یافته و برای انجام این کار نیازی به آموزش مجدد بر روی داده‌های هدف ندارد، که آن را برای کاربردهای در دنیای واقعی بسیار کارآمد و مقیاس‌پذیر می‌سازد.

## ۴-۲ یادگیری پیوسته در مدل‌های بینایی-زبان

بیشتر مدل‌های زبانی بزرگ و بینایی-زبان، در شرایط ایستا آموزش می‌بینند و توانایی کمی در انطباق با داده‌های جدید، بدون بازآموزی کامل، دارند. این محدودیت، باعث توسعه و پیاده‌سازی روش‌های متنوع یادگیری پیوسته در این مدل‌ها شده است. در این زمینه، ژنگ و همکاران [۲۸] طبقه‌بندی شکل ۱-۲ را برای روش‌های متفاوت یادگیری پیوسته ارائه داده‌اند که در ادامه به آن پرداخته می‌شود.

<sup>36</sup>Lightweight decoder





شکل ۲-۱: طبقه‌بندی روش‌های یادگیری پیوسته در مدل‌های بینایی بزرگ و بینایی-زبان [۲۸]

## ۲-۴-۱ روش‌های مبتنی بر حافظه

در این روش، مدل بخشی از داده‌های قدیمی را ذخیره کرده و در کنار داده‌های جدید برای آموزش مجدد استفاده می‌کند. این رویکرد با هدف کاهش پدیده‌ی فراموشی طراحی شده است، که در آن مدل، دانش قبلی خود را هنگام یادگیری اطلاعات جدید از دست می‌دهد. محدودیت در ذخیره‌سازی داده‌های قدیمی چالش اصلی این روش است [۲۸]. گارگ و همکاران [۲۹]، روشی برای آموزش پیوسته‌ی مدل‌های بینایی-زبان مانند CLIP، در مواجهه شدن با داده‌های وب‌مقیاس و در حال تغییر زمانی، ارائه کرده‌اند. این روش با بهره‌گیری از بازپخش داده‌های گذشته و استفاده از مدل پیش‌آمورخته به‌عنوان نقطه شروع، امکان به‌روزرسانی کارآمد مدل را، بدون نیاز به بازآموزی کامل، فراهم می‌کند.

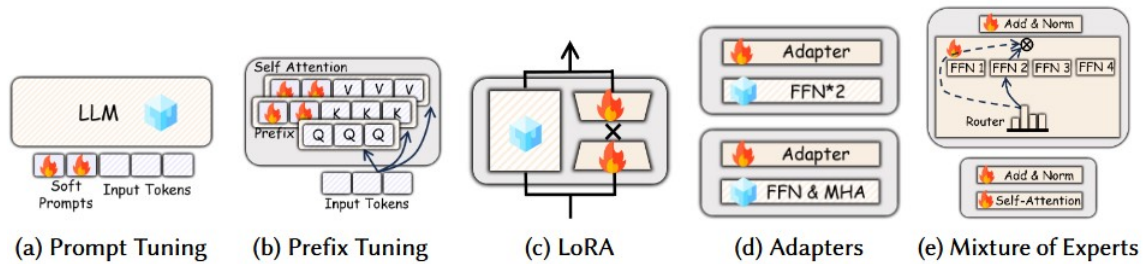
## ۲-۴-۲ روش‌های مبتنی بر تنظیم

در این روش، مدل از مکانیزم‌هایی مانند جریمه یا محدودسازی برای حفظ اطلاعات قبلی استفاده می‌کند. هدف این است که پارامترهایی که در یادگیری گذشته مهم بوده‌اند، هنگام آموزش جدید کمتر تغییر کنند. در این روش، در صورت حجم زیاد وظایف، اثربخشی کاهش می‌یابد [۲۸].

## ۲-۴-۳ روش‌های مبتنی بر تقطیر دانش

مدل دانش خود را از مدل‌های قبلی یا معلم یاد می‌گیرد و توجه به مواردی مانند حساسیت به دقت مدل معلم و انتخاب صحیح داده‌ها برای تقطیر حائز اهمیت است [۲۸]. لو و همکاران [۳۰]، با هدف کاهش فراموشی در یادگیری افزایشی ویدیو، از روشی به نام تقطیر توجه<sup>۳۷</sup> استفاده کرده‌اند که در آن ویژگی‌های توجه از خروجی کدگشای ترنسفورمر CLIP، به مدل جدید منتقل می‌شود. این رویکرد به

<sup>۳۷</sup> Attention distillation



شکل ۲-۲: روش‌های یادگیری پیوسته مبتنی بر معماری در مدل‌های زبانی بزرگ و بینایی-زبان [۲۸]

مدل کمک می‌کند تا دانش مراحل قبلی را حفظ کرده و در عین حال بتواند دسته‌های جدید را بدون نیاز به آموزش کامل مجدد یاد بگیرد.

## ۴-۴-۲ روش‌های مبتنی بر معماری

در این رویکرد، معماری مدل برای جذب وظایف جدید، بدون تداخل با وظایف قبلی، تغییر می‌کند؛ مانند افزودن پیمانه<sup>۳۸</sup> استفاده از تطبیق‌دهنده‌ها و .... ژنگ و همکاران [۲۸]، با بررسی روش‌های مختلف ارائه شده در این رویکرد، طبقه‌بندی مطابق با شکل ۲-۲ ارائه داده اند که در ادامه به بررسی آن‌ها پرداخته می‌شود.

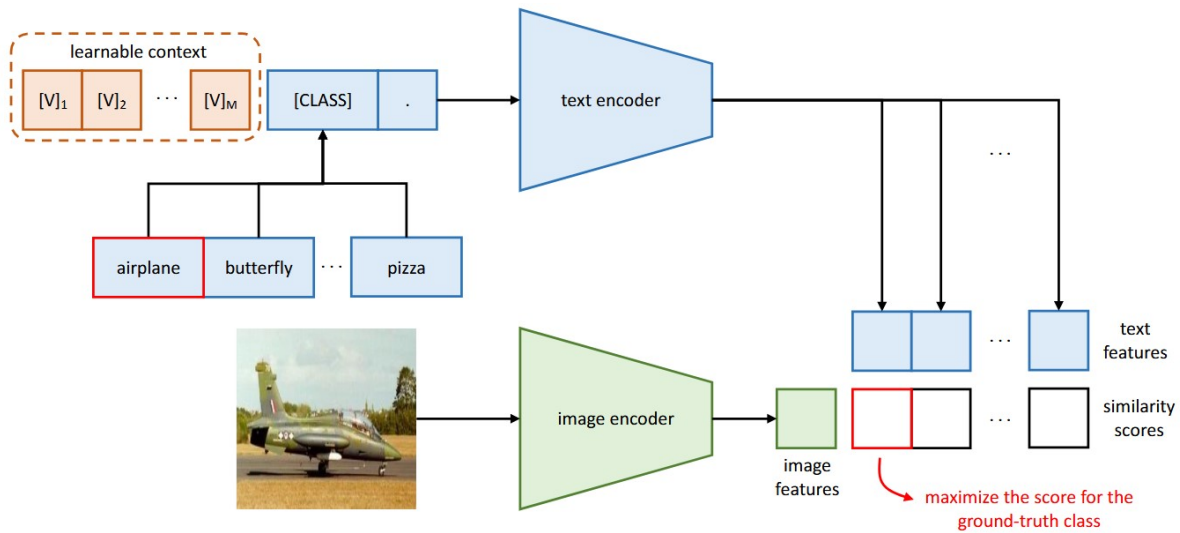
### تنظیم پرامپت

همانطور که در بخش یادگیری انتقالی مدل‌های بینایی-زبان ذکر شد، در روش تنظیم پرامپت، به جای بازآموزی کامل یا تغییر در پارامترهای اصلی مدل، مجموعه‌ای از بردارهای قابل آموزش به عنوان پرامپت به ابتدای نشانه‌های ورودی<sup>۳۹</sup> اضافه می‌شوند. این بردارها بدون دستکاری در ساختار درونی مدل، نقش راهنما را ایفا کرده و جهت‌گیری مدل در تفسیر داده‌های جدید را مشخص می‌سازند. در واقع، مدل با همان دانش قبلی خود به تحلیل ورودی می‌پردازد، اما به واسطه‌ی پرامپت‌های جدید، قادر به تطبیق با وظایف تازه می‌شود [۲۸]. این روش به دلیل مصرف کم منابع محاسباتی و عدم نیاز به تغییر در پارامترهای اصلی، به‌ویژه برای سناریوهایی با دسترسی محدود به مدل یا منابع، بسیار مناسب است و الگوی مورد استفاده در این روش، می‌تواند به خوبی برای مسائل یادگیری پیوسته نیز استفاده شود [۲۸]. ژو و همکاران [۲۲]، رویکردی به نام بهینه‌سازی بافت<sup>۴۰</sup> (که در بخش قبل اشاره شد) ارائه کرده‌اند که

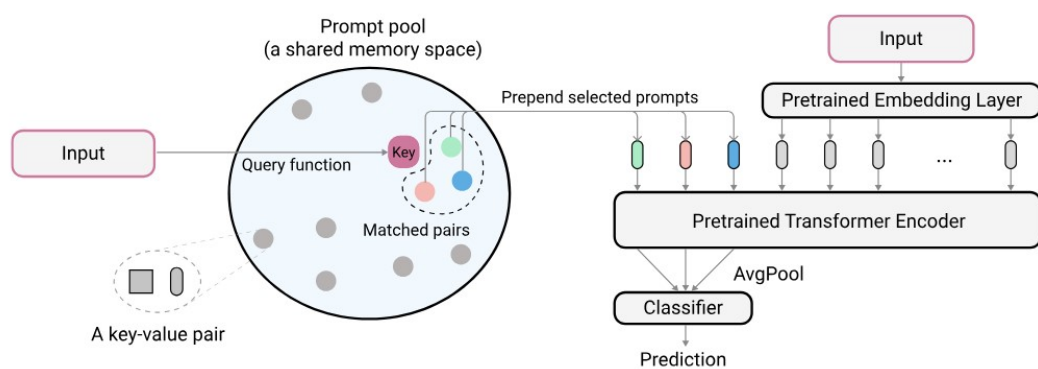
<sup>38</sup>Module

<sup>39</sup>Input tokens

<sup>40</sup>Context Optimization (CoOp)



شکل ۲-۳: نمایش روش CoOp [۲۲]



شکل ۲-۴: نمایش روش L2P در زمان آزمون [۳۱]

به جای پرامپت‌های ثابت، از پرامپت به صورت بردارهای بافت یادگرفتنی<sup>۴۱</sup> که در کنار برچسب متنی داده‌ها قرار می‌گیرند، استفاده می‌کند. مطابق شکل ۲-۳، تمام وزن‌های مدل CLIP، ثابت نگه داشته شده و تنها بردارهای پرامپت، قابل آموزش هستند. به دنبال توسعه ی روش های پرامپت گذاری، روش L2P<sup>۴۲</sup> که توسط وانگ و همکاران [۳۱] ارائه شده است، یک چارچوب نوآورانه برای یادگیری پیوسته بدون نیاز به شناسایی وظیفه در زمان آزمون می‌باشد. همان‌طور که در شکل ۲-۴ مشاهده می‌شود، این روش به جای تغییر وزن‌های مدل پیش‌آمخته، از مجموعه‌ای از پرامپت‌های یادگرفتنی بهره می‌برد که در یک فضای حافظه اشتراکی به نام استخر پرامپت<sup>۴۳</sup> نگهداری می‌شوند. L2P از یک مکانیزم پرس‌وجوی مبتنی بر جفت‌های کلید-مقدار بهره می‌برد تا به صورت پویا و متناسب با ورودی، پرامپت‌های مرتبط را انتخاب کرده و به نشانه‌های ورودی مدل، اضافه کند. سپس این نشانه‌های توسعه‌یافته به مدل پیش‌آمخته تزریق شده و پیش‌بینی انجام می‌شود. در این روش، پرامپت‌ها، دانش خاص هر وظیفه یا دانش مشترک بین وظایف را به صورت فشرده ذخیره می‌کنند و باعث کاهش چشمگیر فراموشی مخرب در یادگیری وظایف متوالی می‌شوند. ساختار طراحی شده در L2P همان‌طور که در شکل ۲-۴ نشان داده شده، از یک بخش انتخاب پرامپت، لایه‌های کدگذار پیش‌آمخته<sup>۴۴</sup>، و دسته‌بند نهایی تشکیل شده است. جهت بهبود و مقاوم‌سازی نسبت به فراموشی در انتخاب پرامپت روش‌های پیشین، مارتین و همکاران [۳۲]، روش STAR-Prompt را معرفی کرده‌اند که از رویکردی دوسطحی برای تنظیم پرامپت، پیروی می‌کند. ابتدا از CLIP، برای تولید پرامپت‌های متنی و ساخت نمونه‌های اولیه<sup>۴۵</sup> پایدار دسته‌ها، استفاده می‌شود و سپس این نمونه‌های اولیه به عنوان کلید برای بازیابی پرامپت‌های تصویری در ترنسفورمر تصویر به کار می‌روند. هم‌چنین روش‌های DualPrompt [۳۳] و H-prompts [۳۴]، مانند مطالعات مذکور، در زمینه‌ی تولید پرامپت‌های مشترک و خاص وظایف ارائه شده‌اند. داهوین و همکاران [۳۵]، از پرامپت اختصاصی برای هر نمونه به جای هر دسته، استفاده کرده‌اند. برخلاف روش‌های پرامپت‌گذاری قبلی، هنگ و همکاران [۳۶]، علاوه بر استفاده از یک پرامپت به جای چندپرامپت، از نمونه‌های پرت مصنوعی برای ایجاد مرز دسته‌بندی بهتر استفاده کرده‌اند. در مطالعات دیگری مانند روش یکپارچه‌سازی دانش بدون تداخل و آگاه از توزیع [۳۷]<sup>۴۶</sup>، به رفع چالش مداخله‌ی پرامپت در تصمیم‌گیری مکانیزم توجه پرداخته شده است. به دنبال پیشرفت‌های روش‌های پرامپت در حوزه‌ی تصویر، ویلا و همکاران [۳۸]، روشی به نام PIVOT

<sup>41</sup> Learnable context vectors<sup>42</sup> Learning to Prompt for Continual Learning<sup>43</sup> Prompt pool<sup>44</sup> Pretrained encoder layers<sup>45</sup> Prototypes<sup>46</sup> Distribution-aware Interference-free Knowledge Integration (DIKI)

را معرفی کرده‌اند که با بهره‌گیری از دانش پیش‌آموخته مدل تصویر-متن CLIP و استفاده از پرامپت‌های مکانی<sup>۴۷</sup> و زمانی، وابستگی‌های زمانی و مکانی ویدیوها را مدل کرده است. وانگ و همکاران [۳۹] نیز با هدف ارتقای عملکرد مدل CLIP در تشخیص حرکات‌های انسانی در ویدیوها، چارچوبی معرفی کرده‌اند که با استفاده از مدل‌سازی حرکتی و پرامپت‌های پویا، به شکلی مؤثر اطلاعات حرکتی را وارد فرآیند یادگیری می‌کند بدون اینکه به تغییر پارامترهای اصلی CLIP نیاز باشد. در مواردی نیز مانند روش ViLT-CLIP [۴۰]، از هر دو نوع پرامپت برای تصویر و متن برای درک ویژگی‌های ویدیویی استفاده شده‌است.

### تنظیم پیشوند

در روش تنظیم پیشوند<sup>۴۸</sup>، مجموعه‌ای از پارامترهای قابل آموزش به عنوان پیشوند به ابتدای هر لایه‌ی ترنسفورمر افزوده می‌شود تا رفتار مدل را در انجام وظایف خاص تنظیم کند. این پیشوندها نقش تغییرات زمینه‌ای را ایفا کرده و برخلاف تنظیم پرامپت، چندین لایه از مدل را تحت تأثیر قرار می‌دهند [۲۸]. روی و همکاران [۴۱]، روشی معرفی کرده‌اند که با استفاده از پیشوندهای قابل یادگیری در هر لایه‌ی مدل، امکان یادگیری وظایف جدید را بدون فراموشی وظایف قبلی فراهم می‌کند. این پیشوندها با ترکیب کانولوشن و اطلاعات مشترک بین وظایف، باعث انتقال بهتر دانش و کاهش تعداد پارامترهای لازم در یادگیری پیوسته می‌شوند.

### سازگاری رتبه پایین

با وارد کردن ماتریس‌های رتبه پایین در لایه‌های معینی از مدل پیش‌آموخته و منجمد، روش سازگاری رتبه پایین<sup>۴۹</sup>، امکان تنظیم هدفمند بخش‌هایی از مدل را بدون بازآموزی کامل فراهم می‌سازد [۲۸]. مارتین و همکاران [۴۲]، به نتیجه رسیدند که در زمینه‌ی یادگیری پیوسته، جایگزینی روش مذکور به جای تنظیم پرامپت، بدون افزایش چشمگیر در تعداد پارامترها، منجر به بهبود عملکرد مدل می‌گردد.

### وفق دهنده

همانطور که در بخش یادگیری انتقالی مدل‌های بینایی-زبان ذکر شد، وفق دهنده‌ها شبکه‌های عصبی کوچک با ساختار فشرده‌ای هستند که بین لایه‌های مدل اصلی قرار می‌گیرند و به مدل امکان می‌دهند

<sup>47</sup>Spatial prompts

<sup>48</sup>Prefix tuning

<sup>49</sup>Low-Rank Adaptation (LoRA)

ویژگی‌های جدید را یاد بگیرد، بدون آن که نیازی به تغییر پارامترهای اصلی از پیش آموزش دیده باشد [۲۸]. به همین دلیل، می‌توان این روش را نیز در یادگیری پیوسته استفاده نمود. در برخی روش‌ها مانند DIA [۴۳] و EASE [۴۴]، از وفق دهنده‌های سبک‌وزن و اختصاصی برای هر وظیفه‌ی جدید استفاده می‌شود تا مدل بتواند بدون بازآموزی کامل یا ذخیره داده‌های قدیمی، دانش جدید را جذب کند. این دو روش، امکان به‌روزرسانی مدل را بدون آسیب به دانش قبلی فراهم کرده و تصمیم‌گیری ترکیبی میان دسته‌های قدیم و جدید را ممکن می‌سازند. دانگ و همکاران [۴۵]، در مطالعه‌ای دیگر، روشی به نام C-ADA، ارائه داده‌اند که با استفاده از وفق دهنده‌هایی با قابلیت گسترش پارامتری، یادگیری وظایف جدید را بدون نیاز به ذخیره داده‌های گذشته ممکن می‌سازد. این روش با حفظ پارامترهای قبلی و افزودن وزن‌های جدید، از تداخل دانش جلوگیری کرده و عملکرد و سرعت آموزش را به‌طور محسوسی بهبود می‌دهد. برای رفع چالش تداخل پارامترهای دسته‌های مشابه در یادگیری پیوسته، هانگ و همکاران [۴۶]، با استفاده از وفق دهنده‌های قابل تنظیم، ابتدا بازنمایی‌های متنی را متناسب با تأثیر دسته‌های جدید بر دسته‌های قدیمی اصلاح می‌کنند و سپس با یک راه‌برد تجزیه و ادغام پارامترها، فراموشی مدل در حین تنظیم وفق دهنده‌ها را کاهش می‌دهند. در حوزه‌ی ویدیو نیز، پن و همکاران [۴۷]، روشی به نام ST-Adapter پیشنهاد داده‌اند که با افزودن قابلیت زمانی-مکانی به مدل تصویری از پیش‌آموخته، آن را برای وظایف ویدیویی قابل استفاده کرده است.

### مخلوط خبره‌ها

روش مخلوط خبره‌ها<sup>۵۰</sup>، با استفاده از یک مکانیزم دروازه‌ای، به‌صورت پویا تعدادی از شبکه‌های عصبی خبره را برای انجام هر وظیفه فعال می‌کند. این ساختار باعث می‌شود مدل بخش‌های مختلف خود را به وظایف متنوع اختصاص دهد و عملکرد بهتر و مقیاس‌پذیری بیشتری پیدا کند. ونگ و همکاران [۴۸]، در این زمینه رویکردی را ارائه داده‌اند که مدل به‌صورت خودکار تصمیم می‌گیرد که بسته به تغییر داده یا وظیفه، از کدام وفق دهنده‌های موجود استفاده کند یا وفق دهنده‌ی جدیدی اضافه نماید، تا تعادلی میان حفظ دانش قبلی و یادگیری دانش جدید ایجاد شود. در ادامه نیز، یو و همکاران [۴۹]، با استفاده از همین روش و با گسترش تدریجی مدل CLIP و استفاده از مسیرهای انتخابی میان وفق دهنده‌های خبره و مدل اصلی، قابلیت تشخیص یادگیری بدون نمونه حفظ شده و در عین حال بار محاسباتی به شکل چشم‌گیری کاهش می‌یابد.

<sup>50</sup>Mixture of Experts (MoE)

## فصل سوم

### نگارش صحیح

## ۱-۳ مقدمه

فصل مقدمه یک پایان نامه، با بیان نیاز موضوع، تعریف مسئله و اهمیت آن در یک یا چند بند (پاراگراف) آغاز می‌شود<sup>۱</sup> و با مرور پیشینه موضوع (سابقه کارهای انجام‌شده پیشین که ارتباط مستقیمی با مسئله مورد بررسی دارند) ادامه می‌یابد. سپس در یک یا دو بند توضیح داده می‌شود که در این پایان نامه، چه دیدگاه یا راهکار جدیدی نسبت به مسئله (موضوع) مورد بررسی وجود دارد. به عبارت دیگر نوآوری‌ها به صورت کاملاً شفاف و صریح بیان می‌شود. در ادامه ممکن است به نتایج بدست‌آمده نیز به طور مختصر و کلی اشاره شود. در آخرین بند از مقدمه به محتوای فصل‌های بعدی پایان نامه به اختصار اشاره می‌شود. برای مشاهده دستورالعمل کامل دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران) به [؟] یا به سایت کتابخانه دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران) مراجعه نمایید.

نگارش صحیح یک پایان نامه در فهم آسان آن بسیار موثر است. در این فصل مهمترین قواعد نگارشی که باید مورد توجه جدی نگارنده قرار گیرد، به اختصار بیان می‌شود. این قواعد را می‌توان در محورهای اصلی زیر دسته‌بندی کرد:

- فارسی نویسی
- رعایت املای صحیح
- رعایت قواعد نشانه‌گذاری

## ۲-۳ فارسی نویسی

در حد امکان سعی کنید به جای کلمات غیرفارسی از معادل فارسی آنها استفاده کنید، به‌ویژه در مواردی که معادل فارسی مصطلح و رایج است. به‌طور مثال استفاده از کلمه «لذا» به جای «برای همین» یا «به‌همین دلیل» توجیهی ندارد. همچنین کلمه «پردازش» زیباتر از «پروسس» و معادل فارسی «ریزپردازنده» مناسب‌تر از «میکروپروسسور» است. در این‌گونه موارد چنانچه احتمال عدم آشنایی خواننده با معادل فارسی وجود دارد، یا اصطلاح غیرفارسی معمول‌تر است، در اولین ظهور کلمه فارسی، اصل غیرفارسی آن به صورت پاورقی آورده شود. اگر به‌ناچار باید کلمات انگلیسی در لابه‌لای جملات گنجانده شوند، از هر طرف یک فاصله بین آنها و کلمات فارسی پیش و پس از آنها در نظر گرفته شود.

<sup>۱</sup> شروع مقدمه نباید چنان طولانی باشد که هدف اصلی را تحت تاثیر قرار دهد.



چنانچه در پایان نامه از مختصرنویسی استفاده شود، لازم است در اولین استفاده، تفصیل آن در پاورقی آورده شود.

### ۳-۳ رعایت املای صحیح

رعایت املای صحیح فارسی به مطالعه و درک راحت تر کمک می کند. همچنین در نوشته های فارسی باید در حد امکان از همزه «ء، ا، و، ه، إ، ئ» استفاده نشود. به عنوان مثال «اجزاء هواپیما» و «آئین نگارش» ناصحیح، اما «اجزای هواپیما» و «آیین نگارش» صحیح هستند.

### ۴-۳ رعایت قواعد نشانه گذاری

منظور از نشانه گذاری به کاربردن علامت ها و نشانه هایی است که خواندن و فهم درست یک جمله را ممکن و آسان می کند. در ادامه نشانه های معمول و متداول در زبان فارسی و موارد کاربرد آنها به اختصار معرفی می شوند.

#### ۱-۴-۳ ویرگول

ویرگول نشانه ضرورت یک مکث کوتاه است و در موارد زیر به کار می رود:

- در میان دو کلمه که احتمال داده شود خواننده آنها را با کسره اضافه بخواند، یا نبودن ویرگول موجب بروز اشتباه در خواندن جمله شود.

- در موردی که کلمه یا عبارتی به عنوان توضیح، در ضمن یک جمله آورده شود. مثلاً برای کنترل وضعیت فضاپیماها، به دلیل آن که در خارج از جو هستند، نمی توان از بالک های آیرودینامیکی استفاده کرد.

- جداکردن بخش های مختلف یک نشانی یا یک مرجع

- موارد دیگر از این قبیل

پیش از ویرگول نباید فاصله گذاشته شود و پس از آن یک فاصله لازم است و بیشتر از آن صحیح نیست.

### ۲-۴-۳ نقطه

نقطه نشانه پایان یک جمله است. پیش از نقطه نباید فاصله گذاشته شود و پس از آن یک فاصله لازم است و بیشتر از آن صحیح نیست.

### ۳-۴-۳ دونقطه

موارد کاربرد دونقطه عبارتند از:

- پیش از نقل قول مستقیم
  - پیش از بیان تفصیل مطلبی که به اجمال به آن اشاره شده است.
  - پس از واژه‌ای که معنی آن در برابرش آورده و نوشته می‌شود.
  - پس از کلمات تفسیرکننده از قبیل «یعنی» و ...
- پیش از دونقطه نباید فاصله گذاشته شود و پس از آن یک فاصله لازم است و بیشتر از آن صحیح نیست.

### ۴-۴-۳ گیومه

موارد کاربرد گیومه عبارتند از:

- وقتی که عین گفته یا نوشته کسی را در ضمن نوشته و مطلب خود می‌آوریم.
- در آغاز و پایان کلمات و اصطلاحات علمی و یا هر کلمه و عبارتی که باید به صورت ممتاز از قسمت‌های دیگر نشان داده شود.
- در ذکر عنوان مقاله‌ها، رساله‌ها، اشعار، روزنامه‌ها و ...

### ۵-۴-۳ نشانه پرسشی

پیش از «؟» نباید فاصله گذاشته شود و پس از آن یک فاصله لازم است و بیشتر از آن صحیح نیست.

### ۳-۴-۶ خط تیره

موارد کاربرد خط تیره عبارتند از:

- جداکردن عبارت‌های توضیحی، بدل، عطف بیان و ...
- به‌جای حرف اضافه «تا» و «به» بین تاریخ‌ها، اعداد و کلمات

### ۳-۴-۷ پرانتز

موارد کاربرد پرانتز عبارتند از:

- به‌معنی «یا» و «یعنی» و وقتی که یک کلمه یا عبارت را برای توضیح بیشتر کلام بیاورند.
  - وقتی که نویسنده بخواهد آگاهی‌های بیشتر (اطلاعات تکمیلی) به خواننده عرضه کند.
  - برای ذکر مرجع در پایان مثال‌ها و شواهد.
- نکته: بین کلمه یا عبارت داخل پرانتز و پرانتز باز و بسته نباید فاصله وجود داشته باشد.

## ۳-۵ جدا یا سرهم نوشتن برخی کلمات

تقریباً تمامی کلمات مرکب در زبان فارسی باید از هم جدا نوشته شوند؛ به استثنای صفات فاعلی مانند «عملگر»، «باغبان» و یا «دانشمند» و کلماتی نظیر «اینکه»، «آنها». در ادامه به نمونه‌هایی از مواردی که باید اجزای یک کلمه جدا، اما بدون فاصله نوشته شوند، اشاره می‌شود:

۱. در افعال مضارع و ماضی استمراری که با «می» شروع می‌شوند، لازم است که در عین جدا نوشتن، «می» از بخش بعدی فعل جدا نیافتد. برای این منظور باید از «فاصله متصل» استفاده و «می» در اول فعل با <sup>۲</sup>SS از آن جدا شود. به‌طور مثال «می‌شود» به‌جای «می شود».

۲. «ها» جمع باید از کلمه جمع بسته‌شده جدا نوشته شود؛ مگر در برخی کلمات مانند «آنها». این امر در مورد کلمات غیرفارسی که وارد زبان فارسی شده‌اند و با حرف «ها» جمع بسته می‌شوند، مانند «کانال‌ها» یا «فرمول‌ها» مورد تاکید است.

<sup>۲</sup>Shift+Ctrl+@

۳. حروف اضافه مانند «به» وقتی به صورت ترکیب ثابت همراه کلمه پس از خود آورده می شوند، بهتر است با SS از آن جدا شوند. مانند «به صورت»، «به عنوان» و «به لحاظ». لازم به ذکر است هنگامی که حرف اضافه «به» با کلمه پس از خود معنای قیدی داشته باشد، مثل «بشدت» یا «بسادگی»، بهتر است که به صورت چسبیده نوشته شود.

۴. کلمات فارسی نباید با قواعد عربی جمع بسته شوند؛ پس «پیشنهادهای» صحیح و «پیشنهادات» اشتباه است.

۵. اسمها و صفت های دو قسمتی مثل «خط چین» و «نوشته شده» با SS از هم جدا می شود.

۶. شناسه ها با SS از کلمه اصلی جدا می شود. مثل «شده اند» و «شده است».

۷. «است» هنگامی که نقش شناسه را داشته باشد توسط SS از قسمت اصلی جدا می شود. مانند «گفته است».

۸. بند پیشین نباید باعث افراط در استفاده از فاصله متصل شود. مثلاً عبارت «نوشته می شود» صحیح و عبارت «نوشته می شود» ناصحیح است.

۹. فعل های دو کلمه ای که معنای اجزای آنها کاملاً با معنای کل متفاوت است، بهتر است که با SS از هم جدا شوند.

۱۰. کلمات مرکب مثل کلمه «دو کلمه ای» در عبارت «فعل های دو کلمه ای» و «یادداشت برداری».

۱۱. مصدر های دو قسمتی با SS از هم جدا می شوند. مثل «ذوب کردن» و «وارد کردن».

۱۲. صفات تفضیلی مثل «آسان تر».

## فصل چهارم

### مشخصات یک پایان نامه و گزارش علمی

اگرچه برای همه انواع نوشته‌ها، مشخصات و ویژگی‌های واحد و معینی نمی‌توان ذکر کرد، با این حال در یک پایان نامه یا گزارش علمی باید نکات و موارد کلی که در این فصل ذکر می‌شود، بطور کامل رعایت شده باشد.

دقت کنید که پس از عنوان فصل باید حداقل توضیحی کوتاه در مورد موضوع نوشته شود و نمی‌توان مستقیماً بعد از آن عنوان بخش را نوشت و همین طور پس از عناوین بخش‌ها و زیربخش‌ها. (مانند دستورالعمل حاضر)

## ۴-۱ برخورداری از غنای علمی

یک پایان نامه باید پیش از هر چیز به لحاظ علمی از غنای لازم برخوردار باشد. یعنی هدف و پیام روشنی داشته باشد و از پیش‌زمینه علمی، بیان دلایل علمی، ارجاعات مورد نیاز و نتیجه‌گیری شفاف بهره ببرد.

## ۴-۲ ارجاع به موقع و صحیح به منابع دیگر

هر جمله‌ای که در یک پایان نامه نوشته می‌شود یا یک جمله کاملاً بدیهی است یا باید دلیل آن بیان شود و یا اینکه باید به منبعی که آن موضوع را نقل یا اثبات کرده، ارجاع داده شود. اگر مطلب یا گفتاری از منبعی عیناً در گزارش نقل می‌شود، باید آن مطلب داخل گیومه قرار گیرد و با ذکر ماخذ و شماره صفحه، به آن اشاره گردد.

## ۴-۳ ساده‌نویسی

سادگی از ضروریات یک نوشته است. نویسنده باید ساده، روان و در عین حال شیوا و رسا بنویسد و عبارات مبهم، جملات پیچیده و کلمات نامأنوس در نوشته خود به کار نبرد. اگر چه افراط در این امر نیز، به شیوایی نوشته صدمه می‌زند. به کارگیری لغات و اصطلاحات دشوار و دور از ذهن و عبارات و جملات نامنظم و مبهم موجب ایجاد اشکال در فهم خواننده خواهد شد.

برای ساده‌نویسی باید در حد امکان از به کارگیری کلمات «می‌بایست»، «بایستی»، «گردید»، «بوده باشد» و مانند آنها که تکلف‌آور، غلط مصطلح و یا غیرشیوا هستند، به جای «باید»، «است»، «شد» و مثل آنها، اجتناب شود. همین طور، «درجهت» نمی‌تواند جایگزین خوبی برای کلمه روانی مثل «برای» باشد. کلمات و جملات روان و ساده می‌توانند اغلب مفاهیم را براحتی منتقل کنند.

دقت در تنظیم بندها (پاراگرافها) نیز کمک شایانی به روانی و سادگی فهم مطلب می‌کند. بندهای طولانی نیز مانند جملات طولانی می‌توانند خسته‌کننده باشند و خواننده را سردرگم کنند. یک بند نباید کمتر از سه یا چهار سطر یا بیشتر از ۱۰ تا ۱۵ سطر باشد.

## ۴-۴ وحدت موضوع

نویسنده باید در سراسر نوشته از اصل موضوع دور نیافتد و تمام بحث‌ها، مثال‌ها و اجزای نوشته با هماهنگی کامل، پیرامون موضوع اصلی باشد و تاثیری واحد در ذهن خواننده القا کند.

## ۴-۵ اختصار

پایان نامه یا گزارش علمی باید در حد امکان، مختصر و مفید باشد و از بحث‌های غیر ضروری در آن پرهیز شود. نوشتن مطالب ارزشمندی که هیچ ربطی به موضوع ندارد، فاقد ارزش علمی است.

## ۴-۶ رعایت نکات دستوری و نشانه‌گذاری

در سراسر پایان نامه باید قواعد دستوری رعایت شود و ارکان و اجزای جمله در جای مناسب خود آورده شود. همچنین رعایت قواعد نشانه‌گذاری سبب می‌شود که بیان نویسنده روشن باشد و خواننده به سهولت و با کمترین صرف انرژی مطالب را مطالعه و درک کند.

## ۴-۷ توجه به معلومات ذهنی مخاطب

نویسنده باید همواره مخاطب خود را در برابر خود تصور کند و با توجه به معلومات ذهنی مخاطب تمامی پیش‌نیازهای لازم برای درک مطالب مورد بحث را، از پیش برای مخاطب فراهم کند.

## ۴-۸ رعایت مراحل اصولی نگارش

هر کار علمی زمانی به بهترین شکل قابل انجام است که بر اساس یک برنامه‌ریزی مشخص انجام شود. تهیه یک متن علمی با کیفیت نیز نیازمند برنامه‌ریزی مناسب و اجرای منظم آن می‌باشد. مراحل نگارش

را عموماً می‌توان به ترتیب زیر در نظر گرفت:

- تهیه فهرستی از عناوین اصلی و فرعی که باید نوشته شود
  - اولویت‌بندی و تعیین ترتیب منطقی فصل‌ها و بخش‌های گزارش
  - گردآوری اطلاعات اولیه راجع به هر بخش و زیربخش
  - تدوین مطالب جدیدی که باید به قلم نگارنده به گزارش اضافه شود
  - تایپ کردن مطالب با رعایت کامل نکاتی که در این دستورالعمل آموزش داده می‌شود
- رعایت نظم و ترتیب در اجرای مراحل ذکر شده هم فرآیند تهیه پایان نامه یا گزارش علمی را برای نگارنده آسان می‌کند و هم کیفیت نگارش را به میزان قابل توجهی افزایش می‌دهد.



## فصل پنجم

### جمع‌بندی و نتیجه‌گیری و پیشنهادات

در پایان گزارش‌های علمی و فنی لازم است که جمع‌بندی یا نتیجه‌گیری نهایی ارائه شود. در این موارد می‌توان آخرین فصل پایان نامه که پیش از مراجع قرار می‌گیرد را به این امر اختصاص داد.

## ۵-۱ پیشنهادات

در این بخش پیشنهاداتی که محقق جهت ادامه تحقیقات دارد ارائه می‌گردد. دقت شود که پیشنهادات باید از تحقیق انجام شده و نتایج آن حاصل شده باشد و از ذکر جملات کلی باید پرهیز کرد.

## فصل ششم

### موارد به‌روزرسانی

## ۱-۶ استفاده از subfigure

برای شکل‌های ۱-۶ و ۲-۶ به کد مراجعه نمایید



test3 (ج)



test2 (ب)



test1 (آ)

شکل ۱-۶: دانشگاه امیرکبیر



test6 (ب)



test5 (آ)

شکل ۲-۶: پلی تکنیک

## ۲-۶ فرض

فرض ۱-۲-۶. برای استفاده از فرض کافیت از فرمت زیر استفاده کنید

`\begin{asum}`

`\end{asum}`

## ۳-۶ جدول

## ۱-۳-۶ چندسطری

مثالی از جدول چند سطری را می‌توانید در ۱-۶ مشاهده نمایید.

جدول ۱-۶: مقایسه الگوریتم‌های بهینه‌ساز

تابع هزینه	تعداد پارامترها	سناریو	
۲/۴۲۴۱	۱۱	۱	GWO
۹/۱۳۸۹	۸	۲	
۶/۳۳۴۳	۷	۳	
۱/۶۱۷۸۸	۴	۴	
۲۰/۵۸۰۱	۱۱	۱	PSO
۱۰/۱۴۰۲	۸	۲	
۳/۸۳۶۵۹	۷	۳	
۴/۸۶۵۲۱	۴	۴	
۳۴/۸۶۲۷	۱۱	۱	GA
۳۳/۷۱۰۹	۸	۲	
۲۶/۴۲۵۷	۷	۳	
۲۵/۶۳۹۲	۴	۴	

## ۲-۳-۶ مقیاس بندی

برای تنظیم جدول می‌توانید از دستور زیر استفاده نمایید

```
\adjustbox{max width=\textwidth,max totalheight=\textheight, keepaspectratio}{
\begin{tabular}...
.
.
.
\end{tabular}}
```

یا

```
\adjustbox{width=\columnwidth,max totalheight=\textheight, keepaspectratio}{
```

\begin{tabular}...

.

\end{tabular}}

نمونه جدول بدون تنظیم

$N$	SW	$d_\varepsilon$	$d_{L_\star}$	$d_{L_\star}$	$d_{L_\star}$	$k$	$d$	$c_2^\gamma$	$c_1^\gamma$	$c_2^\alpha$	$c_1^\alpha$	
۵۸	۱۱۰	۱/۵	۳۵/۹۲۳۷	۲۴/۹۳۶۲	۱۵/۳۶۹۶	۱/۰۷	۸/۶۶۰۳	۱۱/۲	۴/۳	۲/۴	۶/۶	مثلت
۷۲	۷۵	۰/۹	۲۹/۰۴۹۹	۲۰/۶۵۰۹	۱۳/۲۰۷۴	۱/۰۹	۵/۸۷۷۹	۱۲/۷	۴/۹	۲/۸	۶/۱	پنج شلعی
۷۲	۷۰	۱	۲۳/۹۱۶۹	۱۷/۵۶۶۷	۱۱/۲۳۴۹	۱/۱	۵	۱۳/۳	۵/۲	۲/۳	۵/۹	شش ضلعی

جدول ۶-۲: پارامترهای الگوریتم

نمونه جدول با تنظیم سایز

$N$	SW	$d_\varepsilon$	$d_{L_\star}$	$d_{L_\star}$	$d_{L_\star}$	$k$	$d$	$c_2^\gamma$	$c_1^\gamma$	$c_2^\alpha$	$c_1^\alpha$	
۵۸	۱۱۰	۱/۵	۳۵/۹۲۳۷	۲۴/۹۳۶۲	۱۵/۳۶۹۶	۱/۰۷	۸/۶۶۰۳	۱۱/۲	۴/۳	۲/۴	۶/۶	مثلت
۷۲	۷۵	۰/۹	۲۹/۰۴۹۹	۲۰/۶۵۰۹	۱۳/۲۰۷۴	۱/۰۹	۵/۸۷۷۹	۱۲/۷	۴/۹	۲/۸	۶/۱	پنج شلعی
۷۲	۷۰	۱	۲۳/۹۱۶۹	۱۷/۵۶۶۷	۱۱/۲۳۴۹	۱/۱	۵	۱۳/۳	۵/۲	۲/۳	۵/۹	شش ضلعی

جدول ۶-۳: پارامترهای الگوریتم

$N$	SW	$d_\varepsilon$	$d_{L_\star}$	$d_{L_\star}$	$d_{L_\star}$	$k$	$d$	$c_2^\gamma$	$c_1^\gamma$	$c_2^\alpha$	$c_1^\alpha$	
۵۸	۱۱۰	۱/۵	۳۵/۹۲۳۷	۲۴/۹۳۶۲	۱۵/۳۶۹۶	۱/۰۷	۸/۶۶۰۳	۱۱/۲	۴/۳	۲/۴	۶/۶	مثلت
۷۲	۷۵	۰/۹	۲۹/۰۴۹۹	۲۰/۶۵۰۹	۱۳/۲۰۷۴	۱/۰۹	۵/۸۷۷۹	۱۲/۷	۴/۹	۲/۸	۶/۱	پنج شلعی
۷۲	۷۰	۱	۲۳/۹۱۶۹	۱۷/۵۶۶۷	۱۱/۲۳۴۹	۱/۱	۵	۱۳/۳	۵/۲	۲/۳	۵/۹	شش ضلعی

جدول ۶-۴: پارامترهای الگوریتم

### ۳-۳-۶ استفاده از قضیه

قضیه ۳-۶-۱. تست

□

برهان. تست

برای تغییر کلمه برهان به اثبات، می‌توانید خط ۵۲ در commands.tex را غیر فعال نمایید.  
اشکال هر سطح از itemize را می‌توانید با مراجعه به commands.tex تغییر دهید، کامنت‌های مربوطه

برای اینکار مشخص شده است (خطوط ۴۴ تا ۵۰).  
نمونه‌ای از itemize چهار سطحی:

● تست

○ تست

— تست

\* تست

## منابع و مراجع

- [1] Wang, Liyuan, Zhang, Xingxing, Su, Hang, and Zhu, Jun. A comprehensive survey of continual learning: Theory, method and application. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
- [2] Mai, Zheda, Li, Ruiwen, Jeong, Jihwan, Quispe, David, Kim, Hyunwoo, and Sanner, Scott. Online continual learning in image classification: An empirical survey. Neurocomputing, 469:28–51, 2022.
- [3] van de Ven, Gido M. and Tolias, Andreas S. Three scenarios for continual learning, 2019.
- [4] Churamani, Nikhil, Kara, Ozgur, and Gunes, Hatice. Domain-Incremental Continual Learning for Mitigating Bias in Facial Expression and Action Unit Recognition . IEEE Transactions on Affective Computing, 14(04):3191–3206, October 2023.
- [5] Ma, Jiawei, Tao, Xiaoyu, Ma, Jianxing, Hong, Xiaopeng, and Gong, Yihong. Class incremental learning for video action classification. in 2021 IEEE International Conference on Image Processing (ICIP), pp. 504–508, 2021.
- [6] Park, Jaeyoo, Kang, Minsoo, and Han, Bohyung. Class-incremental learning for action recognition in videos. in 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 13678–13687, 2021.
- [7] Kirkpatrick, James, Pascanu, Razvan, Rabinowitz, Neil, Veness, Joel, Desjardins, Guillaume, Rusu, Andrei A., Milan, Kieran, Quan, John, Ramalho, Tiago, Grabska-



- Barwinska, Agnieszka, Hassabis, Demis, Clopath, Claudia, Kumaran, Dharshan, and Hadsell, Raia. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017.
- [8] Li, Zhizhong and Hoiem, Derek. Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40:2935–2947, 2016.
- [9] Shin, Hanul, Lee, Jung Kwon, Kim, Jaehong, and Kim, Jiwon. Continual learning with deep generative replay. in Guyon, I., Luxburg, U. Von, Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., eds. , *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.
- [10] Chaudhry, Arslan, Ranzato, Marc’Aurelio, Rohrbach, Marcus, and Elhoseiny, Mohamed. Efficient lifelong learning with a-gem. *ArXiv*, abs/1812.00420, 2018.
- [11] Rebuffi, Sylvestre-Alvise, Kolesnikov, Alexander, Sperl, Georg, and Lampert, Christoph H. icarl: Incremental classifier and representation learning. in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5533–5542, 2017.
- [12] Aljundi, Rahaf, Caccia, Lucas, Belilovsky, Eugene, Caccia, Massimo, Lin, Min, Charlin, Laurent, and Tuytelaars, Tinne. Online continual learning with maximally interfered retrieval. *ArXiv*, abs/1908.04742, 2019.
- [13] Minhas, Rashid, Mohammed, Abdul Adeel, and Wu, Q. M. Jonathan. Incremental learning in human action recognition based on snippets. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(11):1529–1541, 2012.
- [14] Li, Tianjiao, Ke, Qiuhong, Rahmani, Hossein, Ho, Rui En, Ding, Henghui, and Liu, Jun. Else-net: Elastic semantic network for continual action recognition from skeleton data. in 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 13414–13423, 2021.

- [15] Cheng, Jian, Liu, Haijun, Wang, Feng, Li, Hongsheng, and Zhu, Ce. Silhouette analysis for human action recognition based on supervised temporal t-sne and incremental learning. *IEEE Transactions on Image Processing*, 24(10):3203–3217, 2015.
- [16] Parisi, German I., Tani, Jun, Weber, Cornelius, and Wermter, Stefan. Lifelong learning of human actions with deep neural network self-organization. *Neural Networks*, 96:137–149, 2017.
- [17] Zhang, Jingyi, Huang, Jiaxing, Jin, Sheng, and Lu, Shijian. Vision-language models for vision tasks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(8):5625–5644, 2024.
- [18] Radford, Alec, Kim, Jong Wook, Hallacy, Chris, Ramesh, Aditya, Goh, Gabriel, Agarwal, Sandhini, Sastry, Girish, Askell, Amanda, Mishkin, Pamela, Clark, Jack, Krueger, Gretchen, and Sutskever, Ilya. Learning transferable visual models from natural language supervision. in *International Conference on Machine Learning*, 2021.
- [19] Singh, Amanpreet, Hu, Ronghang, Goswami, Vedanuj, Couairon, Guillaume, Galuba, Wojciech, Rohrbach, Marcus, and Kiela, Douwe. Flava: A foundational language and vision alignment model. in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15617–15629, 2022.
- [20] Li, Liunian Harold, Zhang, Pengchuan, Zhang, Haotian, Yang, Jianwei, Li, Chunyuan, Zhong, Yiwu, Wang, Lijuan, Yuan, Lu, Zhang, Lei, Hwang, Jenq-Neng, Chang, Kai-Wei, and Gao, Jianfeng. Grounded language-image pre-training. in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10955–10965, 2022.
- [21] Wang, Mengmeng, Xing, Jiazheng, and Liu, Yong. Actionclip: A new paradigm for video action recognition. *ArXiv*, abs/2109.08472, 2021.

- [22] Zhou, Kaiyang, Yang, Jingkang, Loy, Chen Change, and Liu, Ziwei. Learning to prompt for vision-language models. *International Journal of Computer Vision*, 130(9):2337–2348, July 2022.
- [23] Gao, Peng, Geng, Shijie, Zhang, Renrui, Ma, Teli, Fang, Rongyao, Zhang, Yongfeng, Li, Hongsheng, and Qiao, Yu. Clip-adapter: Better vision-language models with feature adapters. *International Journal of Computer Vision*, 132(2):581–595, Feb 2024.
- [24] Wortsman, Mitchell, Ilharco, Gabriel, Kim, Jong Wook, Li, Mike, Kornblith, Simon, Roelofs, Rebecca, Lopes, Raphael Gontijo, Hajishirzi, Hannaneh, Farhadi, Ali, Namkoong, Hongseok, and Schmidt, Ludwig. Robust fine-tuning of zero-shot models. in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7949–7961, 2022.
- [25] Weng, Zejia, Yang, Xitong, Li, Ang, Wu, Zuxuan, and Jiang, Yu-Gang. Open-vclip: Transforming clip to an open-vocabulary video model via interpolated weight optimization. in *ICML*, 2023.
- [26] Gu, Xiuye, Lin, Tsung-Yi, Kuo, Weicheng, and Cui, Yin. Open-vocabulary detection via vision and language knowledge distillation. *arXiv preprint arXiv:2104.13921*, 2021.
- [27] Lüddecke, Timo and Ecker, Alexander. Image segmentation using text and image prompts. in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7076–7086, 2022.
- [28] Zheng, Junhao, Qiu, Shengjie, Shi, Chengming, and Ma, Qianli. Towards lifelong learning of large language models: A survey. *ACM Comput. Surv.*, 57(8), March 2025.
- [29] Garg, Saurabh, Farajtabar, Mehrdad, Pouransari, Hadi, Vemulapalli, Raviteja, Mehta, Sachin, Tuzel, Oncel, Shankar, Vaishaal, and Faghri, Fartash. Tic-clip: Continual train-

- ing of clip models. in The Twelfth International Conference on Learning Representations (ICLR), 2024.
- [30] Lu, Shuyun, Jiao, Jian, Wang, Lanxiao, Qiu, Heqian, Lin, Xingtao, Mei, Hefei, and Li, Hongliang. Video class-incremental learning with clip based transformer. in 2024 IEEE International Conference on Image Processing (ICIP), pp. 500–506, 2024.
- [31] Wang, Zifeng, Zhang, Zizhao, Lee, Chen-Yu, Zhang, Han, Sun, Ruoxi, Ren, Xiaoqi, Su, Guolong, Perot, Vincent, Dy, Jennifer, and Pfister, Tomas. Learning to prompt for continual learning. in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 139–149, 2022.
- [32] Menabue, Martin, Frascaroli, Emanuele, Boschini, Matteo, Sangineto, Enver, Bonicelli, Lorenzo, Porrello, Angelo, and Calderara, Simone. Semantic residual prompts for continual learning. in Leonardis, Aleš, Ricci, Elisa, Roth, Stefan, Russakovsky, Olga, Sattler, Torsten, and Varol, Gül, eds. , Computer Vision – ECCV 2024, pp. 1–18, Cham, 2025. Springer Nature Switzerland.
- [33] Wang, Zifeng, Zhang, Zizhao, Ebrahimi, Sayna, Sun, Ruoxi, Zhang, Han, Lee, Chen-Yu, Ren, Xiaoqi, Su, Guolong, Perot, Vincent, Dy, Jennifer, and Pfister, Tomas. Dual-prompt: Complementary prompting for rehearsal-free continual learning. in Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXVI, p. 631–648, Berlin, Heidelberg, 2022. Springer-Verlag.
- [34] Zuo, Yukun, Yao, Hantao, Yu, Lu, Zhuang, Liansheng, and Xu, Changsheng. Hierarchical prompts for rehearsal-free continual learning. ArXiv, abs/2401.11544, 2024.
- [35] Jung, Dahuin, Han, Dongyoon, Bang, Jihwan, and Song, Hwanjun. Generating instance-level prompts for rehearsal-free continual learning. in 2023 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 11813–11823, 2023.

- [36] Huang, Wei-Cheng, Chen, Chun-Fu, and Hsu, Hsiang. OVOR: Oneprompt with virtual outlier regularization for rehearsal-free class-incremental learning. in The Twelfth International Conference on Learning Representations, 2024.
- [37] Tang, Longxiang, Tian, Zhuotao, Li, Kai, He, Chunming, Zhou, Hantao, Zhao, Hengshuang, Li, Xiu, and Jia, Jiaya. Mind the interference: Retaining pre-trained knowledge in parameter efficient continual learning of vision-language models. in Leonardi, Aleš, Ricci, Elisa, Roth, Stefan, Russakovsky, Olga, Sattler, Torsten, and Varol, Gül, eds. , Computer Vision – ECCV 2024, pp. 346–365, Cham, 2024. Springer Nature Switzerland.
- [38] Villa, Andrés, Alcázar, Juan León, Alfarra, Motasem, Alhamoud, Kumail, Hurtado, Julio, Heilbron, Fabian Caba, Soto, Alvaro, and Ghanem, Bernard. Pivot: Prompting for video continual learning. in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 24214–24223, 2023.
- [39] Wang, Qiang, Du, Junlong, Yan, Ke, and Ding, Shouhong. Seeing in flowing: Adapting clip for action recognition with motion prompts learning. in Proceedings of the 31st ACM International Conference on Multimedia, MM '23, p. 5339–5347, New York, NY, USA, 2023. Association for Computing Machinery.
- [40] Wang, Hao, Liu, Fang, Jiao, Licheng, Wang, Jiahao, Hao, Zehua, Li, Shuo, Li, Lingling, Chen, Puhua, and Liu, Xu. Vilt-clip: Video and language tuning clip with multimodal prompt learning and scenario-guided optimization. Proceedings of the AAAI Conference on Artificial Intelligence, 38(6):5390–5400, Mar. 2024.
- [41] Roy, Anurag, Moulick, Riddhiman, Verma, Vinay, Ghosh, Saptarshi, and Das, Abir. Convolutional prompting meets language models for continual learning. in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2024.

- [42] Wistuba, Martin, Teja Sivaprasad, Prabhu, Balles, Lukas, and Zappella, Giovanni. Choice of PEFT Technique in Continual Learning: Prompt Tuning is Not All You Need. arXiv e-prints, p. arXiv:2406.03216, June 2024.
- [43] Li, Jiashuo, Wang, Shaokun, Qian, Bo, He, Yuhang, Wei, Xing, Wang, Qiang, and Gong, Yihong. Dynamic integration of task-specific adapters for class incremental learning, 2025.
- [44] Zhou, Da-Wei, Sun, Hai-Long, Ye, Han-Jia, and Zhan, De-Chuan. Expandable subspace ensemble for pre-trained model-based class-incremental learning. in 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 23554–23564, 2024.
- [45] Gao, Xinyuan, Dong, Songlin, He, Yuhang, Wang, Qiang, and Gong, Yihong. Beyond prompt learning: Continual adapter for efficient rehearsal-free continual learning. in Leonardis, Aleš, Ricci, Elisa, Roth, Stefan, Russakovsky, Olga, Sattler, Torsten, and Varol, Gül, eds. , Computer Vision – ECCV 2024, pp. 89–106, Cham, 2025. Springer Nature Switzerland.
- [46] Huang, Linlan, Cao, Xusheng, Lu, Haori, and Liu, Xialei. Class-incremental learning with clip: Adaptive representation adjustment and parameter fusion. in Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part LIV, p. 214–231, Berlin, Heidelberg, 2024. Springer-Verlag.
- [47] Pan, Junting, Lin, Ziyi, Zhu, Xiatian, Shao, Jing, and Li, Hongsheng. ST-adapter: Parameter-efficient image-to-video transfer learning. in Oh, Alice H., Agarwal, Alekh, Belgrave, Danielle, and Cho, Kyunghyun, eds. , Advances in Neural Information Processing Systems, 2022.
- [48] Wang, Huiyi, Lu, Haodong, Yao, Lina, and Gong, Dong. Self-expansion of pre-trained models with mixture of adapters for continual learning. in NeurIPS 2024 Workshop on Scalable Continual Learning for Lifelong Foundation Models, 2024.

- [49] Yu, Jiazuo, Zhuge, Yunzhi, Zhang, Lu, Hu, Ping, Wang, Dong, Lu, Huchuan, and He, You. Boosting continual learning of vision-language models via mixture-of-experts adapters. pp. 23219–23230, 06 2024.

## پیوست

موضوعات مرتبط با متن گزارش پایان نامه که در یکی از گروه‌های زیر قرار می‌گیرد، در بخش پیوست‌ها آورده شوند:

۱. اثبات‌های ریاضی یا عملیات ریاضی طولانی.

۲. داده و اطلاعات نمونه (های) مورد مطالعه (Case Study) چنانچه طولانی باشد.

۳. نتایج کارهای دیگران چنانچه نیاز به تفصیل باشد.

۴. مجموعه تعاریف متغیرها و پارامترها، چنانچه طولانی بوده و در متن به انجام نرسیده باشد.

## کد میپل

```
with(DifferentialGeometry):  
with(Tensor):  
DGsetup([x, y, z], M)  
frame name: M  
a := evalDG(D_x)  
D_x  
b := evalDG(-2 y z D_x+2 x D_y/z^3-D_z/z^2)
```



# واژه‌نامه‌ی فارسی به انگلیسی

آ	Cartesian product . . . . حاصل ضرب دکارتی
اسکالر . . . . . Scalar	خ
ب	Automorphism . . . . . خودریختی
بالابر . . . . . Lift	د
پ	Degree . . . . . درجه
پایا . . . . . Invariant	ر
ت	microprocessor . . . . . ریزپردازنده
تناظر . . . . . Correspondence	ز
ث	Submodule . . . . . زیرمدول
ثابت‌ساز . . . . . Stabilizer	س
ج	Character . . . . . سرشت
جایگشت . . . . . Permutation	ص
چ	Faithful . . . . . صادقانه
چند جمله‌ای . . . . . Polynomial	ض
ح	

Connected . . . . . همبند	Inner product . . . . . ضرب داخلی
ی	ط
Edge . . . . . یال	Loop . . . . . طوقه
	ظ
	Valency . . . . . ظرفیت
	ع
	Nonadjacency . . . . . عدم مجاورت
	ف
	Vector space . . . . . فضای برداری
	ک
	Complete reducibility . . . . . کاملاً تحویل پذیر
	گی
	Graph . . . . . گراف
	م
	Permutation matrix . . . . . ماتریس جایگشتی
	ن
	Disconnected . . . . . ناهمبند
	و
	Invertible . . . . . وارون پذیر
	ه

# واژه‌نامه‌ی انگلیسی به فارسی

<b>A</b>	Homomorphism . . . . . همریختی
Automorphism . . . . . خودریختی	<b>I</b>
<b>B</b>	Invariant . . . . . پایا
Bijection . . . . . دوسویی	<b>L</b>
<b>C</b>	Lift . . . . . بالابر
Cycle group . . . . . گروه دوری	<b>M</b>
<b>D</b>	Module . . . . . مدول
Degree . . . . . درجه	<b>N</b>
<b>E</b>	Natural map . . . . . نگاشت طبیعی
Edge . . . . . یال	<b>O</b>
<b>F</b>	One to One . . . . . یک به یک
Function . . . . . تابع	<b>P</b>
<b>G</b>	Permutation group . . . . . گروه جایگشتی
Group . . . . . گروه	<b>Q</b>
<b>H</b>	

Quotient graph . . . . .	گراف خارج‌قسمتی	Trivial character . . . . .	سرشت بدیهی
<b>R</b>		<b>U</b>	
Reducible . . . . .	تحویل پذیر	Unique . . . . .	منحصربفرد
<b>S</b>		<b>V</b>	
Sequence . . . . .	دنباله	Vector space . . . . .	فضای برداری
<b>T</b>			

# **Abstract**

This page is accurate translation from Persian abstract into English.

## **Key Words:**

Write a 3 to 5 KeyWords is essential. Example: AUT, M.Sc., Ph. D, ..



**Amirkabir University of Technology**  
**(Tehran Polytechnic)**

**Department of ...**

**M. Sc. Thesis**

**Title of Thesis**

**By**

**Name Surname**

**Supervisor**

**Dr.**

**Advisor**

**Dr.**

**Month & Year**