



(بسمه تعالیٰ)  
دانشگاه آزاد اسلامی  
 واحد تهران غرب

West Tehran Branch, Islamic Azad University

فرم پیشنهاد تحقیق  
پایان‌نامه‌ی کارشناسی ارشد

شماره دانشجویی: ۴۰۳۱۳۵۳۹۰	نام و نام خانوادگی دانشجو: عارف گنجانی ساری
دانشکده: فنی مهندسی	رشته و گرایش: مهندسی کامپیوتر - نرم افزار
آدرس، تلفن ثابت و همراه: تهران، کریمخان، خیابان استاد نجات الهی کوچه وثوق زمانی ۰۹۱۰۰۶۷۹۶۱۴	
استاد مشاور: دکتر نیره زاغری	استاد راهنما: دکتر نیره زاغری

ظرفیت استاد راهنما و استاد مشاور بر اساس بخشنامه 73/28996 مورخ 15/2/93 بررسی گردید:

انتخاب استادی توسط دانشجو بلامانع می باشد

انتخاب استاد راهنما/ مشاور بدلیل تکمیل ظرفیت امکان پذیر نمی باشد

علی منصور  
تاریخ و امضاء

اینجانب عارف گنجانی ساری با توجه به آگاهی از تکمیل ظرفیت استاد راهنما/استاد مشاور تعهد می نمایم که عواقب ناشی از تأخیر در صدور کد شناسایی پایان نامه را پذیرفته و حق هر گونه اعتراض را از خود سلب می نمایم.

عارف گنجانی ساری  
تاریخ و امضاء

این قسمت توسط گروه آموزشی و پژوهش دانشکده تکمیل خواهد شد

	تاریخ دریافت پروپوزال توسط گروه آموزشی
	تاریخ تصویب پروپوزال در شورای تخصصی گروه
	شماره و تاریخ تصویب پروپوزال در شورای پژوهشی دانشکده
	تاریخ اخذ کد
	کد پایان نامه



دانشگاه آزاد اسلامی

واحد تهران غرب  
West Tehran Branch, Islamic Azad University  
فرم پیشنهاد تحقیق  
پایان‌نامه‌ی کارشناسی ارشد

عنوان تحقیق به فارسی: طراحی یک روش تنظیم تطبیقی ضریب  $\beta$  برای بهبود پایداری و کارایی الگوریتم DPO در مدل‌های زبانی کدنویس.

عنوان تحقیق به انگلیسی:

**Designing an Adaptive Beta Adjustment Method to Improve the Stability and Performance of Direct Preference Optimization (DPO) in Code-Oriented Language Models**

دانشکده: فنی مهندسی

گروه تخصصی:

گرایش: نرم افزار

نیمسال اخذ واحد پایان نامه: اول ۱۴۰۵

نام دانشجو: عارف

نام خانوادگی دانشجو: گنجانی ساری

رشته تحصیلی: مهندسی کامپیوتر

نیمسال ورود به مقطع جاری: ۱۴۰۴ - ۱۴۰۳

مشخصات اساتید:

کد ملی	محل تولد	تاریخ تولد	ش ش	نام پدر	پایه	مرتبه علمی	نام و نام خانوادگی	مسئولیت
					۱۰	استادیار	نیره زاغری	راهنما
					۱۰	استادیار	علی منصور	مشاور

تاریخ تصویب در شورای گروه تخصصی:

تایید مدیر گروه رشته تحصیلی:

نام و نام خانوادگی مدیر گروه رشته تحصیلی:

تاریخ ارسال به حوزه پژوهش و فناوری دانشکده:

تاریخ تصویب در شورای پژوهش و فناوری دانشکده:

تایید مدیر پژوهش دانشکده:

تایید کارشناس پژوهشی دانشکده:

توجه: لطفاً این فرم با مساعدة و هدایت استاد راهنما تکمیل شود.

۱- اطلاعات مربوط به دانشجو:

نام: عارف نامخانوادگی: گنجائی ساری شماره دانشجویی ۴۰۳۱۳۵۳۹۰ مقطع: کارشناسی ارشد رشته تحصیلی: مهندسی کامپیوتر گروه تخصصی: کامپیوتر گرایش: نرم افزار نامدانشکده: فنی مهندسی سال ورود به مقطع جاری: ۱۴۰۳ نیمسال ورودی: اول آدرس پستی در تهران میدان ... تلفن ثابت محل سکونت: ۰۲۱... تلفن همراه: ۹۱۵۵۳۷۹۶۱۴ پست الکترونیک: aref.ganjaee1@yahoo.com

## 2- اطلاعات مربوط به استاد راهنما: تذکرات:

- دانشجویان دوره کارشناسی ارشد می‌توانند یک استاد راهنما و حداقل دو استاد مشاور و دانشجویان دوره دکتری حداقل تا دو استاد راهنما و دو استاد مشاور انتخاب نمایند.
- در صورتی که استاد راهنما و مشاور مدعو می‌باشد، لازم است سوابق تحصیلی، آموزشی و پژوهشی کامل ایشان (رزومه کامل) شامل فهرست پایان‌نامه‌های کارشناسی ارشد و رساله‌های دکتری دفاع شده و یا در حال انجام که استاد مدعو، راهنمایی و یا مشاوره آنرا بر عهده داشته‌اند، به همراه مدارک مربوطه و همچنین آخرین حکم کارگزینی (حکم هیأت علمی) ضمیمه گردد.
- استاد راهنما و مشاور موظف هستند قبل از پذیرش پرپوزال، به سقف ظرفیت پذیرش خود توجه نموده و در صورت تکمیل بودن ظرفیت پذیرش، از ارسال آن به دانشکده و حوزه پژوهشی و یا در نوبت قراردادن و ایجاد وقفه در کار دانشجویان جداً پرهیز نمایند. بدیهی است در صورت عدم رعایت موازین مربوطه، مسئولیت تأخیر در ارائه پرپوزال و عواقب کار، متوجه گروه تخصصی و دانشکده خواهد بود.

## اطلاعات مربوط به استاد راهنما:

نام و نام خانوادگی: ..... آخرین مدرک تحصیلی دانشگاهی: .....

تخصص اصلی: ..... رتبه دانشگاهی (مرتبه علمی): ..... تلفن همراه: .....

تلفن منزل یا محل کار: ..... نام و نام خانوادگی به زبان انگلیسی: .....

نحوه همکاری با واحد تهران غرب:

مدعو  نیمه وقت  تمام وقت

## اطلاعات مربوط به استاد مشاور:

نام و نام خانوادگی: ..... آخرین مدرک تحصیلی دانشگاهی: .....

تخصص اصلی: ..... رتبه دانشگاهی (مرتبه علمی): ..... تلفن همراه: .....

تلفن منزل یا محل کار: ..... نام و نام خانوادگی به زبان انگلیسی:

نحوه همکاری با واحد تهران غرب:

تمام وقت  مدعو  نیمه وقت

#### 4- اطلاعات مربوط به پایاننامه:

الف- عنوان تحقیق

1- عنوان به زبان فارسی:

طراحی یک روش تنظیم تطبیقی ضریب  $\beta$  برای بهبود پایداری و کارایی الگوریتم DPO در مدل‌های زبانی کدنویس.

2- عنوان به زبان انگلیسی:

Designing an Adaptive Beta Adjustment Method to Improve the Stability and Performance of Direct Preference Optimization (DPO) in Code-Oriented Language Models

ب - تعداد واحد پایاننامه: ۶

ج- بیان مسئله اساسی تحقیق به طور کلی (شامل تشریح مسئله و معرفی آن، بیان جنبه‌های مجهول و مبهم، بیان متغیرهای مربوطه و منظور از تحقیق) :

رشد سریع مدل‌های زبانی بزرگ و کاربرد گسترده آن‌ها در تولید و تکمیل کد باعث شده است که نیاز به روش‌های پایدار، قابل اعتماد و کارآمد در فرآیند آموزش این مدل‌ها بیش از گذشته احساس شود. الگوریتم Direct Preference Optimization (DPO) به عنوان یک جایگزین سبکتر برای RLHF مطرح شده، اما عملکرد آن بهشت وابسته به ضریب  $\beta$  است؛ پارامتری که در صورت انتخاب نامناسب می‌تواند موجب نوسان، عدم پایداری، افزایش بیش از حد KL divergence می‌شود. در حالی‌که مدل‌های کدنویس نیازمند دقت بالا، ثبات در رفتار و توانایی پاسخ‌دهی قابل اتکا هستند، بسیاری از پژوهش‌های موجود  $\beta$  را مقدار ثابت در نظر گرفته‌اند و روش شناخته‌شده‌ای برای تنظیم پویا و داده‌محور این پارامتر ارائه نشده است. نبود مکانیزم تطبیقی برای  $\beta$  سبب می‌شود مدل در مواجهه با داده‌های پیچیده حوزه کدنویسی یا سناریوهای ناهمگن بهینه‌سازی، عملکردی غیرقابل پیش‌بینی از خود نشان دهد. از این‌رو، طراحی یک روش تنظیم تطبیقی ضریب  $\beta$  می‌تواند سهم مهمی در افزایش پایداری، بهبود رفتار گرadiانی، کنترل KL divergence زبانی کرد. کاهش ریسک همگرایی ناپایدار و ارتقای کیفیت نهایی مدل‌های زبانی کدنویس داشته باشد. این موضوع هم از نظر علمی و هم از نظر کاربردی در سامانه‌های کمکی بر نامه‌نویسی، اهمیت ویژه‌ای دارد و می‌تواند زمینه‌ساز توسعه روش‌های کارآمدتر در حوزه Alignment مدل‌های زبانی باشد.

د - اهمیت و ضرورت انجام تحقیق (شامل اختلاف نظرها و خلاصهای تحقیقاتی موجود، میزان نیاز به موضوع، فواید احتمالی نظری و عملی آن و همچنین مواد، روش و یا فرآیند تحقیقی احتمالاً جدیدی که در این تحقیق مورد استفاده قرار می‌گیرد:

گسترش کاربرد مدل‌های زبانی بزرگ در تولید، تکمیل و تحلیل کد، نیاز به روش‌های آموزشی قابل اعتماد و پایدار را بیش از گذشته بر جسته کرده است. هرگونه ناپایداری در فرآیند بهینه‌سازی می‌تواند منجر به

خروجی‌های ناقص، ناهمانه‌نگ یا غیرقابل استفاده در سناریوهای عملی برنامه‌نویسی شود؛ از این‌رو پایداری در آموزش Code LLM‌ها یکی از نیازهای اساسی پژوهش‌های نوین هوش مصنوعی است. الگوریتم DPO به دلیل سادگی و حذف نیاز به مدل پاداش، به عنوان روشی کارآمد برای یادگیری مبتنی بر ترجیحات شناخته می‌شود؛ اما ثابت بودن ضریب  $\beta$  در این الگوریتم باعث می‌شود مدل نسبت به تغییرات داده، ساختار پیچیده کد و شرایط آموزشی مختلف حساسیت بالایی داشته باشد. این مشکل می‌تواند منجر به افزایش خطای KL، نوسان در گرادیان‌ها، همگرایی نامطمئن و افت کیفیت خروجی مدل شود. با وجود تلاش‌هایی برای بهبود DPO، همچنان هیچ سازوکار عملی و داده‌محوری برای تنظیم پویا و هوشمند  $\beta$  ارائه نشده است. این خلاً پژوهشی نه تنها کارایی مدل‌های زبانی کدنویس را محدود می‌کند، بلکه مانع توسعه ابزارهای قابل‌اتکای برنامه‌نویسی شده و موجب افزایش ریسک خروجی‌های نادرست می‌شود. انجام این تحقیق می‌تواند با ارائه یک روش تنظیم تطبیقی  $\beta$ ، به بهبود پایداری فرایند آموزش، کنترل رفتار KL، افزایش کیفیت خروجی مدل و توسعه روش‌های کارآمدتر در زمینه Alignment مدل‌های زبانی کمک کند. به همین دلیل، این موضوع از نظر علمی و کاربردی دارای اهمیت ویژه است.

۵- مرور ادبیات و سوابق مربوطه (بیان مختصر پیشینه تحقیقات انجام شده در داخل و خارج کشور پیرامون موضوع تحقیق و نتایج آنها و مرور ادبیات و چارچوب نظری تحقیق):

در سال‌های اخیر، آموزش مدل‌های زبانی بزرگ بر پایه داده‌های ترجیحی با هدف بهبود کیفیت و همسویی با رفتار انسانی مورد توجه گسترشده قرار گرفته است. الگوریتم Direct Preference Optimization (DPO) به عنوان یک روش ساده‌تر و بدون نیاز به مدل پاداش، جایگزینی مؤثر برای RLHF معرفی شد. پژوهش‌های اولیه نشان می‌دادند که DPO می‌تواند فرایند آموزش را تسهیل کند و خروجی‌هایی با کیفیت قابل قبول ارائه دهد، اما مطالعات بعدی نشان دادند که این الگوریتم در برابر تغییرات داده و تنظیمات آموزشی حساس است. در ادامه، روش‌های مختلفی برای رفع محدودیت‌های DPO پیشنهاد شد. از جمله می‌توان به Cal-DPO اشاره کرد که مسئله تنظیم دما و توزیع احتمال را بررسی می‌کند، یا SafeDPO که باز چینی ترجیحات به دنبال افزایش ایمنی پاسخ‌هایست. همچنین رویکردهای دیگری همچون Contrastive PO، ROPO، Pre-DPO، Unified-PO، PO، SGDPO، GRPO و DPO را کاهش داده و دقت یا پایداری فرایند آموزش را افزایش دهند. با وجود این تلاش‌ها، ضعف‌های DPO را کاهش داده و دقت یا پایداری فرایند آموزش را افزایش دهند. با آنکه  $\beta$  پارامتری تعیین‌کننده در شدت بروزرسانی مدل و رفتار گرادیانی است، تاکنون پژوهشی جامع برای طراحی یک مکانیزم تطبیقی و زمان‌مند جهت تنظیم خودکار این ضریب ارائه نشده است. این خلاً پژوهشی، به‌ویژه در کاربردهایی مانند تولید که پایداری مدل نقش کلیدی دارد، کاملاً محسوس است. بنابراین مطالعه حاضر با هدف پر کردن این شکاف، به دنبال ارائه روشی برای تنظیم تطبیقی  $\beta$  در الگوریتم DPO است تا پایداری، کارایی و کیفیت نهایی مدل‌های زبانی کدنویس ارتقا یابد.

و - جنبه جدید بودن و نوآوری در تحقیق:

نوآوری اصلی این تحقیق در ارائه یک سازوکار پویا و داده‌محور برای تنظیم ضریب  $\beta$  در الگوریتم DPO است؛ پارامتری که در تمام پژوهش‌های پیشین به صورت ثابت انتخاب شده و یکی از عوامل اصلی ناپایداری و افت کارایی این الگوریتم محسوب می‌شود. در این پژوهش،  $\beta$  بر اساس رفتار لحظه‌ای مدل در طول آموزش (مانند تغییرات  $\Delta$ ، KL و شاخص‌های پایداری) به صورت تطبیقی تنظیم می‌شود. این رویکرد، نخستین تلاش ساختارمند برای طراحی یک روش کنترل هوشمند  $\beta$  در زمینه آموزش ترجیح‌محور است و می‌تواند محدودیت‌های شناخته شده DPO را برطرف کند، پایداری و کیفیت یادگیری را افزایش دهد و به‌طور ویژه در مدل‌های زبانی کدنویس که حساسیت بیشتری دارند، کارایی بالاتری ایجاد کند.

ز- اهداف مشخص تحقیق (شامل اهداف آرمانی، کلی، اهداف ویژه و کاربردی): طراحی و ارزیابی یک روش تنظیم تطبیقی ضریب  $\beta$  بهمنظور افزایش پایداری و بهبود کارایی الگوریتم DPO در آموزش مدل‌های زبانی کدنویس.

ح- در صورت داشتن هدف کاربردی، نام بهرهوران (سازمان‌ها، صنایع و یا گروه ذینفعان) ذکر شود (به عبارت دیگر محل اجرای مطالعه موردی):

ط- سوالات تحقیق:

- 1) ضریب  $\beta$  ثابت در الگوریتم DPO چگونه بر پایداری، رفتار KL و کیفیت خروجی مدل‌های زبانی کدنویس تأثیر می‌گذارد؟
- 2) آیا می‌توان یک مکانیزم تطبیقی برای تنظیم پویا و هوشمند  $\beta$  طراحی کرد که در طول فرآیند یادگیری، پایداری و کارایی الگوریتم DPO را بهبود دهد؟
- 3) روش تنظیم تطبیقی  $\beta$  چه میزان بر عملکرد نهایی مدل‌های زبانی کدنویس در مقایسه با DPO کلاسیک مؤثر است؟

ی- فرضیه‌های تحقیق:

استفاده از ضریب  $\beta$  ثابت در الگوریتم DPO موجب ناپایداری و افزایش خطای KL در فرآیند آموزش مدل‌های زبانی کدنویس می‌شود.

تنظیم تطبیقی و پویا ضریب  $\beta$  می‌تواند پایداری و کارایی الگوریتم DPO را نسبت به حالت کلاسیک بهبود دهد.

روش تنظیم تطبیقی  $\beta$  منجر به افزایش کیفیت خروجی مدل‌های زبانی کدنویس در مقایسه با DPO با  $\beta$  ثابت خواهد شد.

ک- تعریف واژه‌ها و اصطلاحات فنی و تخصصی (به صورت مفهومی و عملیاتی):  
الگوریتم (DPO) Direct Preference Optimization

تعریف مفهومی:

DPO یک الگوریتم آموزش مدل‌های زبانی مبتنی بر داده‌های ترجیحی است که بدون نیاز به مدل پاداش، از اختلاف احتمالی میان پاسخ‌های برنده و بازنده برای یادگیری مستقیم استفاده می‌کند.

تعریف عملیاتی:

در این تحقیق، DPO به عنوان الگوریتم پایه انتخاب شده و رفتار آن در شرایط  $\beta$  ثابت و  $\beta$  تطبیقی تحلیل، مقایسه و ارزیابی می‌شود.

ضریب  $\beta$

تعریف مفهومی:

$\beta$  یک پارامتر مقیاس‌دهنده در DPO است که شدت تأثیر اختلاف بین پاسخ ترجیح‌داده شده و پاسخ ردشده را در بهروزرسانی مدل کنترل می‌کند.

تعریف عملیاتی:

در این پژوهش،  $\beta$  به صورت پویا و تطبیقی تنظیم می‌شود و یکی از متغیرهای اصلی در سنجش پایداری و کیفیت آموزش مدل است.

تعریف مفهومی:

KL شاخصی برای سنجش میزان فاصله یا انحراف توزیع احتمالی مدل جدید نسبت به مدل مرجع است و در یادگیری ترجیح محور برای کنترل رفتار مدل استفاده می‌شود.

تعریف عملیاتی:

در این تحقیق، KL یکی از معیارهای اصلی برای ارزیابی پایداری مدل در طول آموزش و یکی از ورودی‌های تنظیم  $\beta$  تطبیقی خواهد بود.

مدل‌های زبانی کدنویس (Code LLMs)

تعریف مفهومی:

Code LLM‌ها مدل‌های زبانی بزرگی هستند که برای تولید، تکمیل یا تحلیل کد برنامه‌نویسی آموزش دیده‌اند و خروجی‌های دقیق و ساختارمند تولید می‌کنند.

تعریف عملیاتی:

در این پژوهش، Code LLM‌ها به عنوان بستر آزمایشی مورد استفاده قرار می‌گیرند تا اثر  $\beta$  تطبیقی بر کیفیت تولید کد و پایداری فرایند آموزش بررسی شود.

تنظیم تطبیقی ( $\beta$  Adaptive Beta Adjustment)

تعریف مفهومی:

تنظیم تطبیقی  $\beta$  روشی است که در آن مقدار  $\beta$  بر اساس رفتار لحظه‌ای مدل، داده‌ها و شاخص‌های پایداری به صورت پویا تغییر می‌کند.

تعریف عملیاتی:

در این پایان‌نامه، مکانیزمی طراحی و پیاده‌سازی می‌شود که مقدار  $\beta$  را در هر مرحله آموزش بر اساس معیارهایی مانند  $\Delta$ ، KL و رفتار گرادیانی تنظیم می‌کند.

داده‌های ترجیحی (Preference Data)

تعریف مفهومی:

نوعی داده شامل دو پاسخ برای یک ورودی است که یکی توسط انسان یا مدل برنده و دیگری بازنده فرض شده است.

تعریف عملیاتی:

در این پژوهش، داده‌های ترجیحی برای آموزش و آزمایش DPO در حالت  $\beta$  ثابت و  $\beta$  تطبیقی مورد استفاده قرار می‌گیرند.

## 5- روش شناسی تحقیق:

الف- شرح کامل روش تحقیق بر حسب هدف، نوع داده‌ها و نحوه اجراء (شامل مواد، تجهیزات و استانداردهای مورد استفاده در قالب مراحل اجرایی تحقیق به تفکیک):

تذکر: در خصوص تفکیک مراحل اجرایی تحقیق و توضیح آن، از به کار بردن عناوین کلی نظیر، «گردآوری اطلاعات اولیه»، «تهیه نمونه‌های آزمون»، «انجام آزمایش‌ها» و غیره خودداری شده و لازم است در هر

مورد توضیحات کامل در رابطه با منابع و مراکز تهیه داده‌ها و ملزومات، نوع فعالیت، مواد، روش‌ها، استانداردها، تجهیزات و مشخصات هر یک ارائه گردد.

این تحقیق از نوع توسعه‌ای-کاربردی است و بر پایه روش‌های پژوهش محاسباتی و تحلیل الگوریتمی انجام می‌شود. هدف اصلی، طراحی و ارزیابی یک سازوکار تطبیقی برای تنظیم ضریب  $\beta$  در الگوریتم DPO است. روش اجرا شامل سه مرحله کلی است:

(1) تحلیل رفتار DPO با  $\beta$  ثابت:

در این مرحله با استفاده از مجموعه‌ای از داده‌های ترجیحی و یک مدل زبانی پایه، رفتار الگوریتم DPO در شرایط مختلف بررسی می‌شود. شاخص‌هایی مانند KL divergence، پایداری گرادیان و کیفیت خروجی مدل تحلیل می‌گردد تا نقاط ضعف وضعیت فعلی مشخص شود.

(2) طراحی مکانیزم تنظیم تطبیقی  $\beta$ :

در این گام، یک روش داده‌محور و پویا برای بهروزرسانی  $\beta$  پیشنهاد می‌شود. این روش بر اساس متغیرهایی مانند اختلاف ترجیح ( $\Delta$ )، میزان انحراف مدل از مدل مرجع (KL)، و رفتار گرادیانی تعریف می‌شود. این مکانیزم در قالب یک الگوریتم قابل پیاده‌سازی تدوین خواهد شد.

(3) ارزیابی روش پیشنهادی:

روش تطبیقی  $\beta$  بر روی یک مدل زبانی کدنویس آزمایشی اعمال شده و نتایج آن با DPO کلاسیک مقایسه می‌شود. معیارهای ارزیابی شامل پایداری آموزشی، تغییرات KL، و کیفیت خروجی مدل در وظایف تولید کد خواهد بود.

در تمامی مراحل از داده‌های ترجیحی استاندارد، محیط‌های محاسباتی رایج و کتابخانه‌های یادگیری عمیق بهره گرفته می‌شود. اجرای آزمایش‌ها در محیط‌های محاسباتی سبک انجام شده و نیازی به تجهیزات آزمایشگاهی فیزیکی وجود ندارد.

ب- متغیرهای مورد بررسی در قالب یک مدل مفهومی و شرح چگونگی بررسی و اندازه‌گیری متغیرها:

متغیر مستقل:

- روش تنظیم ضریب  $\beta$  (ثابت یا تطبیقی)

متغیرهای وابسته:

- میزان پایداری الگوریتم DPO

- مقدار KL divergence در طول آموزش

- کیفیت خروجی مدل زبانی در تولید کد

متغیرهای کنترلی:

- نوع مدل زبانی پایه

- نوع داده ترجیحی

- تنظیمات مشترک آموزش

نحوه اندازه‌گیری:

- پایداری با بررسی تغییرات رفتار گرادیان و نوسانات loss

- KL divergence از طریق مقایسه توزیع مدل و مدل مرجع

- کیفیت خروجی کد با معیارهای متداول ارزیابی و ظایف برنامه‌نویسی

ج- شرح کامل روش (میدانی، کتابخانه‌ای) و ابزار (مشاهده و آزمون، پرسشنامه، مصاحبه، فیش‌برداری و غیره) گردآوری داده‌ها:

این پژوهش از داده‌های ثانویه استفاده می‌کند. داده‌ها شامل مجموعه‌های از نمونه‌های ترجیحی (ورودی، پاسخ برنده و بازنده) هستند که در منابع متن‌باز موجود بوده و برای آموزش الگوریتم‌های alignment می‌شوند.

ابزار گردآوری داده شامل:

- استخراج داده‌های ترجیحی از مخازن استاندارد

- آمده‌سازی داده برای ورودی به الگوریتم DPO

- استفاده از ابزارهای محاسباتی و کتابخانه‌های یادگیری عمیق برای پیده‌سازی و تحلیل

این تحقیق فاقد ابزارهای میدانی مانند پرسشنامه یا مصاحبه است و بهصورت کاملاً محاسباتی انجام می‌شود.

د - جامعه آماری، روش نمونه‌گیری و حجم نمونه (در صورت وجود و امکان):

این تحقیق از نوع محاسباتی و الگوریتمی است و جامعه آماری انسانی ندارد. داده‌های ترجیحی بهصورت مجموعه‌ای

از نمونه‌ها (ورودی + پاسخ برنده/بازنده) مورد استفاده قرار می‌گیرد و نیازی به نمونه‌گیری آماری نیست. "نمونه‌ها"

همان داده‌های ترجیحی هستند که به عنوان ورودی مدل بهکار می‌روند.

ه - روش‌ها و ابزار تجزیه و تحلیل داده‌ها:

تحلیل داده‌ها بهصورت محاسباتی و شامل موارد زیر است:

- تحلیل KL divergence برای بررسی میزان انحراف مدل

- تحلیل پایداری الگوریتم از طریق بررسی رفتار loss و گرادیان

- مقایسه خروجی مدل در حالت  $\beta$  ثابت و  $\beta$  تطبیقی

- استفاده از معیارهای ارزیابی کیفیت تولید کد

ابزار تحلیل شامل محیط‌های برنامه‌نویسی، کتابخانه‌های یادگیری عمیق و ابزارهای محاسباتی استاندارد است. نتایج بهصورت نمودار، جدول و تحلیل‌های کمی گزارش خواهد شد.

## 6- فهرست منابع و مأخذ

فهرست منابع و مأخذمی باشد براساس روش APA باشد.

7- استفاده از امکانات آزمایشگاهی واحد:

آیا برای انجام تحقیقات نیاز به استفاده از امکانات آزمایشگاهی واحد تهران غرب می‌باشد  بله  خیر

در صورت نیاز به امکانات آزمایشگاهی لازم است نوع آزمایشگاه، تجهیزات، مواد و وسایل مورد نیاز در این قسمت مشخص گردد.

نوع آزمایشگاه	تجهیزات مورد نیاز	مواد و وسایل	مقدار مورد نیاز

امضاء مدیرگروه تخصصی:

امضاء استاد راهنما:

8- زمان بندی انجام تحقیق:

الف- تاریخ شروع:.....ج- تاریخ  
ب- مدت زمان انجام تحقیق:..... تمام:.....

تذکر: لازم است کلیه فعالیت‌ها و مراحل اجرایی تحقیق (شامل زمان ارائه گزارشات دوره‌ای) و مدت زمان مورد نیاز برای هر یک، به تفکیک پیش‌بینی و در جدول مربوطه درج گردیده و در هنگام انجام عملی تحقیق، حتی الامکان رعایت گردد.

پیش‌بینی زمان‌بندی فعالیت‌ها و مراحل اجرایی تحقیق و ارائه گزارش پیشرفت کار

زمان اجرا به ماه							زمان کل (ماه)	شرح فعالیت	ردیف
6	5	4	3	2	1				
						1	Unified-PO، Cal-DPO، SafeDPO، مطالعه	1	
							مرور ادبیات مرتبط	2	
							تعریف مسئله، متغیرها و چارچوب نظری	3	
							طراحی اولیه مکانیزم $\beta$	4	
						1	جمع‌آوری داده‌های ترجیحی	5	
							آمده‌سازی داده‌ها	6	
							پیاده‌سازی نسخه پایه DPO	7	
							اجرای تست اولیه و استخراج رفتار KL و پایداری	8	
						1	طراحی نسخه اولیه روش	9	
							انتخاب معیارهای کنترل (Loss behavior، $\Delta$ ، KL)	10	
							پیاده‌سازی اولیه مکانیزم $\beta$	11	
							اجرای تست‌های اولیه و اصلاح مدل	12	
						1	اجرای آزمایش نهایی $\beta$	13	
							مقایسه KL، پایداری، کیفیت خروجی	14	
							استخراج نمودار و جداول	15	
							تحلیل رفتار مدل در چند سناریویی کدنویسی	16	

							تحلیل نهایی نتایج	17
						1	بررسی نکات قوت و محدودیت‌ها	18
							تدوین بخش نتیجه‌گیری و پیشنهادات	19
							نگارش فصول پایان‌نامه	20
						1	تهیه جداول، نمودارها و ضمائم	21
							ارسال نسخه پیش‌نویس به استاد	22
							انجام اصلاحات	23
							آماده‌سازی نسخه نهایی جهت دفاع	24

a

- توجه: 1- زمان و نوع فعالیت‌های اجرایی پایان‌نامه، حتی‌الامکان باید با مندرجات جدول منطبق باشد.  
 2- حداقل زمان قابل قبول برای پیش‌بینی مراحل مطالعاتی و اجرایی پایان‌نامه کارشناسی ارشد 6 ماه و حداقل 12 ماه می‌باشد.

## 9- صور تجلیه گروه تخصصی

## نام و نامخانوادگی دانشجو: نام و نامخانوادگی دانشجو:

نام و نامخانوادگی استاد راهنمای  
امضاء \_\_\_\_\_ تاریخ \_\_\_\_\_ -1

نام و نامخانوادگی استاد مشاور  
امضاء  
تاریخ

نام و نامخانوادگی عضو کمیته نظارت بر تحقیق  
امضاء تاریخ -1

شورای گروه تخصصی ..... در محل ..... در تاریخ ..... با حضور اعضای مربوطه .....  
با ..... — ..... آقای ..... عنوان ..... و ..... موضوع ..... پایان نامه ..... خانم ..... تشکیل .....

## بررسی و به تصویب رسید.

نام و نام خانوادگی اعضای شورای گروه	امضاء	تاریخ
-1		
-2		
-3		
-4		
-5		
-6		

نام و نامخانوادگی مدیرگروه: امضاء تاریخ

## 10- صور تجلیسه شورای یژوهش و فناوری دانشکده:

موضوع و طرح تحقیق پایان نامه دانشجو ..... مقطع کارشناسی ارشد گروه .....  
 رایش ..... با عنوان ..... ک .....  
 در تاریخ ..... به تصویب کمیته گروه تخصصی مربوطه رسیده است، در جلسه .....  
 شماره ..... مورخ ..... شورای پژوهش دانشکده طرح شد و پس از بحث و  
 تبادل نظر مورد تصویب اکثریت اعضاء قرار گرفت.

ردیف	نام و نام خانوادگی	نوع رأی (موافق یا مخالف)	محل امضاء	توضیحات
1				
2				
3				
4				
5				
6				
7				

نام و نام خانوادگی مدیر پژوهش دانشکده:

امضاء .....  
 تاریخ .....

نام و نام خانوادگی رئیس دانشکده:

امضاء .....  
 تاریخ .....

## این فرم باید توسط دانشجو تکمیل شود

فرم سازمان مرکزی دانشگاه آزاد اسلامی

### فرم الف- فرم اطلاعات پایاننامه کارشناسی ارشد

نام واحد دانشگاهی: واحد تهران غرب	
عنوان پایاننامه کارشناسی ارشد:	
نیمسال تحصیلی:	نام و نامخانوادگی دانشجو:
تعداد واحد	شماره دانشجویی: پایاننامه:
کد	گرایش:
علم پایه <input type="checkbox"/>	علوم سانی <input type="checkbox"/> هنر <input type="checkbox"/>
رشته	رشته تحصیلی: رشته:
رشته	نام و نامخانوادگی استاد راهنما:
کد شناسایی استاد راهنما:	مرتبه علمی: استادیا <input type="checkbox"/> استاد <input type="checkbox"/> دانش <input type="checkbox"/>
رشته	نام و نامخانوادگی استاد مشاور:
کد شناسایی استاد مشاور:	مرتبه علمی: استاد <input type="checkbox"/> استاد <input type="checkbox"/> دانش <input type="checkbox"/> مربي <input type="checkbox"/>

فرم تعهد اساتید در قبال نتایج حاصل از پایان نامه کارشناسی ارشد یا رساله دکتری

اینجانبان اساتید راهنمای مشاور و پایاننامه، خانم/آقای  
قطعه دانشجوی، پایاننامه، خانم/آقای  
گرایش رشته .....  
عنوان با .....»

نظر گرفتن حقوق دانشگاه آزاد اسلامی واحد تهران غرب نسبت به انتشار نتایج حاصل از تحقیق مذکور (در قالب کتاب، مقاله، طرح تحقیقاتی، اختراع، اکتشاف و ...) اقدام ننماییم.

## نام و نامخانوادگی استاد راهنما: تاریخ و امضاء

## نام و نامخانوادگی استاد مشاور اول: تاریخ و امضاء

در تاریخ ..... فرم مذبور که به امضاء استادی محترم راهنمای مشاور رسیده است، دریافت گردید.

امضاء کارشناس پژوهش دانشکده:

فرم تعهد دانشجویان در قبال نتایج حاصل از پایاننامه کارشناسی ارشد یا رساله دکتری

اینجانب	.....	دانشجوی	.....	وروדי	.....
.....	.....	.....	.....	.....	.....
.....	.....	گرایش	.....	که موضوع	.....
پایاننامه‌ام	.....	تحت	.....	عنوان:	.....
»	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....

..... در شورای گروه «..... در شورای گروه .....» تخصصی به تصویب رسیده، متعهدمی‌گردم (الف) کلیه مطالب و مندرجات پایاننامه‌ام بر اساس اصول علمی و حاصل از تحقیقات خودم تهیه شود و در صورت استفاده از مطالب، نتایج تحقیقات، نقل قول‌ها، جداول و نمودارهای دیگران در پایاننامه/رساله، منابع و مأخذ آن به نحوی که قابل تشخیص و تفکیک از متن اصلی باشد قید گردد.

ب) در صورتیکه از نتایج تحقیقاتم علاوه بر پایان نامه / رساله، کتاب ، مقاله، اختراع، اکتشاف و هر گونه تولیدات علمی حاصل شود، صرفاً به نام دانشگاه آزاد اسلامی واحد تهران غرب بوده و این موضوع صراحتاً در مکاتبات و تولیدات اینجانب درج و بر اساس ضوابط دانشگاه اقدام نمایم.

ج) در صورت استفاده از کمک‌های مالی و غیر مالی نهادهای دولتی و غیر دولتی از موضوع تحقیق اینجانب مراتب را کتاباً به دانشکده اطلاع دهم در غیر اینصورت دانشکده مجاز به تغییر عنوان پایاننامه یا سایر اقدامات حقوقی می‌باشد.

نام و نامخانوادگی دانشجو  
تاریخ و امضاء

در تاریخ ..... فرم مزبور که توسط آقای / خانم ..... به امضاء ..... رسیده است، دریافت گردید.

امضاء کارشناس پژوهش دانشکده:

فرم شماره 5

اینجانب اساتید راهنما و مشاور پایان نامه / رساله خانم / آقای ..... دانشجوی مقطع ..... رشته ..... گرایش ..... متعهد می‌شویم در کلیه تولیدات علمی مستخرج از این پایان نامه / رساله (اعم از مقاله، طرح پژوهشی، کتاب، اختراع و ....) که به منظور استفاده از تسهیلات تشویقی، سفرهای علمی، فرصت مطالعاتی و امتیاز ارتقاء علمی (اعضاء هیات علمی واحد و ....) ارائه می‌گردد، نسبت به درج آدرس دقیق دانشگاه و واحد و رعایت ترتیب و توالی آن به عنوان تنها آدرس خود بصورت زیر اقدام نمایم. بدیهی است چنانچه تحت هر شرایطی و در هر زمان، دانشگاه خلاف موارد ذکر شده را مشاهده نماید نسبت به تصمیم اتخاذ شده هیچگونه آدعاً نداشته و حق هر گونه اعتراضی را از خود سلب و ساقط می‌نمایم.

آدرس واحد دانشگاهی به فارسی:

نام گروه تخصصی، نام دانشکده، واحد تهران غرب، دانشگاه آزاد اسلامی، تهران، ایران.

آدرس واحد دانشگاهی به انگلیسی:

Department of College of Islamic Azad University, Tehran, Iran

نام و نام خانوادگی استاد راهنما:

تاریخ و امضاء:

نام و نام خانوادگی استاد مشاور:

تاریخ و امضاء:

تصریه:

تکمیل این فرم برای اساتید تمام وقت واحد تهران غرب الزامی است و اساتید نیمهوقت یا مدعو در صورتی که نام ایشان به عنوان نویسنده اول و عهدهدار مکاتبات در تولیدات علمی مستخرج از پایان نامه قید نشود ملزم به امضاء این فرم نمی‌باشند.

اینجانب ..... دانشجوی ورودی ..... مقطع ..... رشته ..... گرایش ..... متعهد می‌شوم در کلیه تولیدات علمی مستخرج از پایان نامه / رساله ام ..... اعزام ..... وان .....

».....

به منظور تسویه حساب و یا استفاده از تسهیلات تشویقی و سفرهای علمی و ..... نسبت به درج آدرس دقیق دانشگاه واحد و رعایت ترتیب و توالی آن بصورت ذیل به عنوان تنها آدرس خود اقدام نمایم. بدیهی است چنانچه تحت هر شرایطی و در هر زمان، دانشگاه خلاف موارد ذکر شده را مشاهده نماید نسبت به تصمیم اتخاذ شده هیچگونه ادعایی نداشته و حق هر گونه اعتراضی را از خود سلب و ساقط می‌نمایم.

Department of College of Islamic Azad University, Tehran, Iran

نام و نام خانوادگی دانشجو:

تاریخ و امضاء:



## فرم تعهد نامه رعایت منشور و موازین اخلاق پژوهش

اینجانب عارف گنجائی ساری دانشجوی ورودی ۱۴۰۳ مقطع کارشناسی ارشد رشته مهندسی کامپیوتر گرایش نرم افزار که موضوع پایان نامه ام تحت عنوان: « طراحی یک روش تنظیم تطبیقی ضریب  $\beta$  برای بهبود پایداری و کارایی الگوریتم DPO در مدل های زبانی کدنویس.» در شورای گروه تخصصی به تصویب رسیده، بدین وسیله معهده می شوم:

۱) ضمن آگاهی و اطلاع کامل از منشور و موازین اخلاق پژوهش مدون کمیته اخلاق پژوهش وزارت علوم، تحقیقات و فناوری، ملزم به رعایت تمامی اصول و حدود اخلاقی اجزای پژوهش می باشم و عواقب ناشی از بروز هرگونه اتفاقات و تخلفات آگاهانه یا ناآگاهانه در حوزه پژوهش را می پذیرم.

۲) کلیه مطالب و مندرجات پایان نامه ام بر اساس اصول علمی و حاصل از تحقیقات خودم (توأم با ویژگی های اصالت، صداقت، بی طرفی، امانت داری، مشورت و نقیضی بر، پایین دی به جستجوی حقیقت و عنصر زمان، دقت در تدوین گزارش پژوهش) تهیه شود و در پژوهش خود مرتکب هیچ گونه رفتار سوء پژوهشی مدون در منشور و موازین اخلاق پژوهش (یکی از مصادیق جعل داده ها، تحریف داده ها، سرقت علمی، اجاره علمی و تعارض منافع) نشوم.

۳) با اطلاع به اینکه درج نام اساتید راهنما و/یا مشاور در مقاله هی مستخرج از پایان نامه/رساله الزامی است، در صورتی که از نتایج تحقیقات علاوه بر پایان نامه/رساله، کتاب، مقاله، اختراع، اکتشاف و هر گونه تولیدات علمی حاصل شود، ضمن اطلاع رسانی به اساتید راهنما و/یا مشاور جهت انتشار با رعایت مالکیت معنوی صرفاً با ضوابط آدرس دهی مشخص شده به نام دانشگاه آزاد اسلامی واحد تهران غرب اقدام نمایم.

۴) چنانچه در هر مقطع زمانی خلاف موارد فوق ثابت شود، عواقب ناشی از آن را می پذیرم و واحد دانشگاهی مجاز است با اینجانب مطابق با ضوابط و مقررات رفتار نموده و به استناد ماده (11) و (12) آیین نامه اجرایی قانون پیشگیری و مقابله با تقلب در تهیه آثار علمی به شماره ۱۳۹۸/۰۵/۳۰ ت ۵۵۶۸۶/۶۶۶۸۲ مورخ ۱۳۹۸/۰۵/۳۰، در صورت ابطال مدرک تحصیلی ام و متعاقباً اطلاع رسانی به محل کار جهت لغو اثر مدرک تحصیلی ام، هیچ گونه ادعایی نخواهم داشت.

عارف گنجائی ساری  
تاریخ و امضاء