

بهمن ماه ۱۴۰۱
نسخه ۱

راهنمای مسابقه جایزه سالانه هوش مصنوعی ایران iAAA

فهرست مطالب

هدف.....	۱-۱
دامنه کاربرد و مخاطبین.....	۱-۲
زمان بندی مسابقه.....	۱-۳
موضوع مسابقه و ماهیت داده.....	۱-۴
تشکیل تیمها.....	۱-۵
پرو فایل افراد و تیمها.....	۱-۶
نحوه دسترسی به داده های مسابقه.....	۱-۷
آموزش مدل.....	۱-۸
کد خروجی.....	۱-۹
کتابخانه های مجاز.....	۱-۱۰
فرمت خروجی کد.....	۱-۱۱
ارسال فایل حاوی کد.....	۱-۱۲
نتایج میان دوره ای.....	۱-۱۳
جدول نتایج تیمها.....	۱-۱۴
ارسال نتایج نهایی.....	۱-۱۵
داوری سیستمی.....	۱-۱۶
داوری حضوری.....	۱-۱۷
جدول زمانی.....	۱-۱۸

۱- هدف

هدف از تنظیم این دستورالعمل، تدوین راهنمایی یکپارچه به منظور برگزاری و شرکت در مسابقه رقابت-محور iAAA است.

۲- دامنه کاربرد و مخاطبین

دستورالعمل حاضر در سطح شرکت‌کنندگان، داوران و تیم اجرایی دبیرخانه جایزه سالانه هوش مصنوعی ایران (iAAA) کاربرد دارد.

۳- زمان‌بندی مسابقه

- رونمایی از مسئله: ۱۳ بهمن ۱۴۰۱
- تشکیل تیم‌ها: ۱ اسفند ۱۴۰۱
- شروع مسابقه: ۴ اسفند ۱۴۰۱
- بازه ارسال کدهای دوره‌ای: از ۱۵ اسفند ۱۴۰۱ تا ۵ اردیبهشت ۱۴۰۲
- اعلام نتایج داوری سیستمی: ۱۰ اردیبهشت ۱۴۰۲
- اختتامیه، اهدای جوایز و رویداد حضوری: ۱۳ اردیبهشت ۱۴۰۲

۴- موضوع مسابقه و ماهیت داده

- موضوع رقابت حاضر، **بررسی خودکار نیاز به جراحی اندو در مدالیته OPG** با استفاده از شرح حال و شرایط بیمار است.
- در این مسئله از شرکت‌کنندگان خواسته می‌شود که با استفاده از عکس OPG و هوش مصنوعی، نیاز/عدم نیاز بیماران به جراحی اندو را پیش‌بینی کنند.
- داده‌ها توسط دندانپزشکان مجرب علامت (Label) زده شده‌اند و فرمت دیتای ورودی تصاویر DICOM است.

۵- تشکیل تیم‌ها

شرکت‌کنندگان اعم از اینکه به صورت فردی یا تیمی ثبت‌نام کرده باشند (یا در فرایند تیم‌سازی اعضای تیم خود را انتخاب کرده باشند)، لازم است **حداکثر تا تاریخ ۱۴۰۱/۱۲/۱ نسبت به ثبت نهایی اطلاعات اعضای تیم خود** از طریق لینک اقدام نمایند.

<https://survey.porsline.ir/s/kjYG1zc>

شرکت‌کنندگان محترم در نظر داشته باشند تعداد اعضای مورد قبول برای هر تیم **حداقل ۱ و حداکثر ۶ نفر** می‌باشد.

لازم به ذکر است حضور یک شرکت‌کننده در دو تیم متمایز جایز نیست و در صورت مغایرت بین اطلاعات ثبت شده در لینک فوق و اطلاع ثبت شده در زمان ثبت‌نام، نسخه آخر (لینک حاضر) مدنظر قرار خواهد گرفت. اطلاعات موردنیاز دبیرخانه از هریک از تیم‌های شرکت‌کننده در مسابقه شامل موارد زیر خواهد بود:

- نام تیم
- عکس لوگوی تیم (نشانی تصویری بیانگر هویت بصری تیم)
- تعداد اعضای تیم
- نام و نام خانوادگی سرپرست تیم و دیگر اعضا
- کد ملی سرپرست تیم و دیگر اعضا (شایان‌ ذکر است در وارد کردن کد ملی سرپرست تیم نهایت دقت را داشته باشید؛ چراکه این آیتم شناسه تیم شما در روند داوری خواهد بود).

۶- پروفایل افراد و تیم‌ها

به‌منظور ارائه پروفایل روان‌شناسی به شرکت‌کنندگان محترم، تمامی افراد شرکت‌کننده که تعامل دارند این اطلاعات را دریافت نمایند و تاکنون اقدام به انجام تست‌های روان‌شناسی ننموده‌اند، می‌توانند از طریق لینک: <https://yek.link/profile> نسبت به انجام تست‌ها اقدام نمایند.

همچنین شرکت‌کنندگانی که **کلید** تست‌های روان‌شناسی را تکمیل ننموده‌اند، می‌توانند از طریق لینک زیر وارد پلتفرم پروفایل شوند:

- <https://iaaa.ai/profile-login>
- کد ملی: User Name
- شماره موبایل بدون صفر اول: Password

پس از ورود، شما می‌توانید نسبت به تکمیل یا اصلاح برخی از اطلاعات خود (و تغییر رمز عبور) اقدام نمایید. برخی موارد توسط دبیرخانه اجرایی و بر اساس نتایج تست‌ها و آزمون‌های شما، ثبت می‌گردد که امکان ویرایش آن توسط شما وجود ندارد. درمورد تعدادی از اطلاعات شخصی، تصمیم شما مبنی بر تعامل به نمایش یا عدم نمایش آن‌ها در پروفایل، محفوظ است.

به‌منظور اطلاع و مشاهده پروفایل کلیه افرادی که در این فرایند شرکت نموده‌اند، می‌توانید به لینک زیر مراجعه کنید (لازم به ذکر است لینک زیر فقط پس از لاگین اطلاعات ذکر شده را به شما نمایش خواهد داد):

<https://iaaa.ai/members>

یادآوری می‌شود امکان جست و جوی کاربران بر اساس برخی فیلدها مانند نام شهر محل سکونت، نام و نام خانوادگی، رشته تحصیلی و دوره های گذرانده شده وجود دارد. همچنین، می‌توانید در پلتفرم فوق نسبت به ارسال پیام، ایجاد گروه (<https://iaaa.ai/groups>)، چت گروهی و... اقدام نمایید.

بدیهی است ثبت دقیق و کامل اطلاعات افراد در کل این فرایند الزامی است.

۷- نحوه دسترسی به داده‌های مسابقه

داده‌های مسابقه از دو طریق قابل دسترسی شرکت‌کنندگان هستند:

• داده موجود بر روی سرورهای iAAA

○ شرکت‌کنندگان از طریق لینک <http://iaaa.ai/data> می‌توانند داده‌ها را دانلود نمایند تا بر روی رایانه‌ها یا سرورهای شخصی و یا سرورهای google colab به ایجاد و آموزش مدل‌ها، پردازند.

داده‌ها در زمان مقرر، در لینک فوق در دسترس قرار خواهند گرفت و در گروه تلگرام، به اطلاع شرکت‌کنندگان خواهد رسید.

• داده موجود بر روی سرورهای مجموعه ابررایانه سیمرغ

○ بدین منظور لازم است شرکت‌کنندگان در سایت <https://simorgh.cloud> با کد تخفیف **iaaa-gpu01** ثبت‌نام نموده و هزینه سرور موردنظر خود را پرداخت نمایند. با مذاکرات انجام شده توسط تیم اجرایی iAAA با مجموعه فوق، امکان استفاده از GPU Server با تخفیف برای شرکت‌کنندگان iAAA جهت ایجاد و آموزش مدل‌های مدنظر، در حال آماده سازی است.

جزئیات و زمان فعال شدن کد تخفیف پس از تنظیمات سرور ها توسط آن مجموعه، در گروه تلگرام، به اطلاع شرکت‌کنندگان خواهد رسید.

۸- آموزش مدل

شما در این مرحله مجاز هستید از هر مدل و کتابخانه‌ای که می‌خواهید استفاده کنید.

۹- کد خروجی

به منظور محاسبه خطا توسط سیستم، شما باید یک **کد py** به دبیرخانه تحویل دهید (نحوه ارسال در ادامه توضیح داده می‌شود) که بتواند به عنوان آرگومان خط فرمان، آدرس پوشه داده‌ها و همچنین اسم فایل خروجی را دریافت کند. سپس، برای هر یک از تصاویر موجود در پوشه، خروجی مدل را به صورت ه یا ا محاسبه کند. دقت کنید که انجام تمام مراحل اعم از خواندن عکس،

پیش‌پردازش، استفاده از مدل برای تولید خروجی، پس‌پردازش و ... باید در این فایل انجام شود. همچنین، باید یک خروجی به فرمت CSV تولید شود که ستون اول آن نام تصاویر و ستون دوم آن ه یا ا است.

به منظور اطمینان از اینکه کد شما به درستی کار کند، تعداد محدودی تصویر به همراه یک کد کوچک در اختیار شما قرار خواهد گرفت تا بتواند کد شما را چک کند. کد شما باید از خط فرمان به صورت زیر قابل اجرا باشد:

```
$ python Submission.py --inputs /path/to/images/folder --output predict.csv
```

برای دریافت خروجی خط فرمان و استفاده راحت‌تر از آن، می‌توانید از قطعه کد نمونه استفاده کنید و داخل تابع پیش‌بینی محاسبه خروجی را انجام دهید. لطفاً دقت کنید هیچ اقدام خاصی در این کد انجام نشده است، ورودی این تابع دو استرینگ است که اولی آدرس پوشه‌ها و دومی آدرس و نام فایل خروجی است.

نکته: کد شما بر روی سرور ممکن است بر روی CPU یا GPU اجرا شود؛ لذا باید چک کنید که کد شما هم بر روی GPU و هم CPU بدون خطا، کار کند (به خصوص در صورتی که از پای‌تورچ استفاده می‌کنید).

۱۰- کتابخانه‌های مجاز

شما در کد محاسبه خروجی خود، مجاز به استفاده از تعداد محدودی کتابخانه هستید. این کتابخانه‌ها در لیست زیر آمده‌اند. تمام این کتابخانه‌ها در آخرین نسخه خود خواهند بود.

1. GDCM
2. GDM
3. Numpy
4. Opencv
5. Pandas
6. Pillow
7. Matplotlib
8. Pydicom
9. Pylibjpg
10. Pylibjpg-libjpg
11. Scikit-learn
12. Tensorflow

- 13. Torch**
- 14. Torchvision**
- 15. Tqdm**
- 16. Keras**
- 17. Scipy**
- 18. plotly**
- 19. Seaborn**
- 20. Albumentation**
- 21. SimpleITK**
- 22. Tensorflow-addons**

نکته: در صورت نیاز به کتابخانه دیگری، با آی دی تلگرام پشتیبان هماهنگ کنید. در صورتی که تشخیص داده شود کتابخانه مورد تقاضا یک کتابخانه ضروری برای حل مسئله خواهد بود به لیست کتابخانه‌های قابل قبول اضافه خواهد شد و این امر تنها با اطلاع رسانی رسمی در گروه تلگرام، نهایی تلقی می‌شود.

۱۱- فرمت خروجی کد

خروجی مطلوب برای فایل CSV باید فرمت مناسبی داشته باشد، در غیر این صورت کد تصحیح نمی‌تواند به درستی دقت کد شما را محاسبه کند و حتی ممکن است در این روند دچار خطا شود. دقت کنید که فایل CSV شما باید فقط دارای دو ستون باشد که ستون اول نام فایل ورودی (بدون نام پوشه‌ای که تصویر در آن قرار دارد) و ستون دوم مقدار عددی ۱ یا ۰ باشد که ۱ به معنی وجود بیماری و ۰ به معنی عدم وجود بیماری است.

همچنین در این فایل نباید عبارتی برای سرستون‌ها درج شود. به عنوان مثال، یک نمونه صحیح از این خروجی را می‌توانید در زیر مشاهده کنید:

```
Open  [v] [⌂]
```

```
1 1.jpg,1|
2 10.jpg,0
3 100.jpg,1
4 1000.jpg,0
5 1001.jpg,1
6 1002.jpg,1
7 1003.jpg,1
8 1004.jpg,1
9 1005.jpg,1
10 1006.jpg,0
11 1007.jpg,0
12 1008.jpg,0
13 1009.jpg,0
14 101.jpg,1
```

۱۲- ارسال فایل حاوی کد

همان‌طور که درمورد فرمت خروجی کد پیش‌بینی ذکر شد، ارسال فایل باید دقیقاً طبق فرمت مشخصی اتفاق بیفتد. در غیر این صورت، ممکن است در فرایند تصحیح خطایی رخ بدهد.

داخل فایل ارسالی شما باید یک کد **Submission.py** باشد (نام این کد باید دقیقاً همین عبارت باشد و نام دیگری قابل قبول نیست). این کد (Submission.py) مدل و داده را خوانده و نتایج را به صورت فایل CSV ایجاد می‌نماید. همچنین، شما می‌توانید در کنار این فایل، چندین فایل دیگر و فایل مدل را قرار دهید.

- فایل ارسالی تنها با فرمت **zip** قابل قبول است (فرمت‌های rar, 7z, gz غیرقابل قبول هستند).
- نام فایل ارسالی باید به صورت **TeamID_#.zip** باشد؛ آی دی تیم که همان کد ملی سرپرست/سرگروه تیم است و سپس آندرلاین و بعد نوبت ارسال نتایج به صورت دو رقمی همانند زیر:

0082582889_05.zip

- نوبت ارسال نتایج: هر تیم می‌تواند در بازه زمانی مشخص شده **تا حداکثر ۳۰ بار** مدل خود را ارسال نماید.
- حجم فایل ارسالی **نباید بیشتر از ۱۵۰ مگابایت** باشد.
- جهت ارسال فایل، سرپرست تیم باید از لینک زیر اقدام نماید. اطلاعات کاربری جهت ورود، به ایمیل ایشان ارسال می‌گردد (در صورت وجود هرگونه مشکل پس از تاریخ ۵ اسفندماه، به آی دی پشتیبان، پیام داده شود).

<https://upload.iAAA.ai>

۱۳- نتایج میان‌دوره‌ای

تیم‌ها می‌توانند در بازه ۱۵ اسفند ۱۴۰۱ تا ۵ اردیبهشت ۱۴۰۲ نسبت به ارسال نتایج دوره‌ای اقدام نمایند. در ارسال نتایج لازم است به موارد زیر، توجه شود:

- تشابه مدل‌ها و کدهای تیم‌های متمایز بررسی شده و در صورت احراز تشابه هر دو تیم از روند مسابقه حذف می‌شوند (معیارهای موردنظر به‌منظور احراز تشابه؛ پیش‌پردازش، ساخت معماری مدل، آموزش و پس‌پردازش خواهند بود).
- کدها باید Reproducible باشند.
- تعداد دفعات مجاز ارسال نتایج برای هر تیم حداکثر ۳۰ مرتبه است که بر اساس ID تیم (کد ملی سرپرست تیم) ردیابی خواهد شد.
- کلیه کدهای ارسال شده در هر مرتبه و نسخه آن‌ها توسط تیم، نگهداری و آرشیو شود (PreProcess, Train, PostProcess Submission, ...).
- ارسال نتایج روی سرور مربوطه از طریق مراجعه به آدرس <http://upload.iAAA.ai> و ورود به‌وسیله نام کاربری و رمز عبور سرپرست هر تیم ممکن خواهد بود.

۱۴- جدول نتایج تیم‌ها

دسترسی به جدول نتایج میان‌دوره‌ای از طریق leaderboard لینک زیر خواهد بود.

<https://iAAA.ai/leaderboard>

لازم به ذکر است لینک فوق همان پلتفرم درج شده در بخش ۶- **پرو فایل افراد و تیم‌ها** است؛ لذا فارغ از علاقه‌مندی به دریافت پرو فایل روان‌شناسی (حتی در صورت عدم علاقه به این موضوع و

عدم تکمیل تست‌های روان‌شناسی)، لازم است کلیه افراد از طریق لینک زیر در این پلتفرم ثبت‌نام نمایند:

<https://iaaa.ai/registerbp>

۱۵- ارسال نتایج نهایی

پس از ۵ اردیبهشت ۱۴۰۲ دیگر امکان ارسال نتایج میان‌دوره‌ای برای تیم‌ها وجود نداشته و تیم‌ها باید تا ۲ روز بعد (یعنی ۱۴۰۲/۲/۷) تنها ۲ مدل از میان کلیه مدل‌های ارسال شده خود را انتخاب و ارسال نمایند (این انتخاب حسب نظر تیم بوده و می‌تواند بر اساس نتایج کسب شده در leaderboard باشد). این ۲ مدل در واقع مدل‌های برتر هر تیم، برای مسابقه iAAA است.

۱۶- داوری سیستمی

نتیجه اجرای ۲ کد برتر کلیه تیم‌ها، بر اساس نتیجه اجرای فرایند داوری سیستمی و بر اساس پلتفرم تصحیح کدهای مسابقه iAAA سطح‌بندی شده و ۶ تیم برتر انتخاب می‌گردند. نکات زیر در فرایند داوری سیستمی، یادآوری می‌شوند:

- تعیین و توضیح معیار تصحیح سیستم بر اساس طبقه‌بندی دو کلاسه است (F1 SCORE).
- سایر پارامترهایی که گزارش می‌گردند:
 - زمان پردازش
 - حجم مدل
- ۶ تیم برتر در سایت مسابقه برای دعوت به مرحله دفاع و داوری حضوری معرفی می‌شوند.
- ۲ کد برتر ارسال شده توسط کلیه تیم‌ها توسط دبیرخانه iAAA اجرا شده از میان ۲ کد ارسال شده، کد برتر و با نتایج بهتر از هر تیم، توسط دبیرخانه انتخاب می‌شود.
- سنجش مدل‌ها روی ۳۰ درصد دیتای تست در جدول leaderboard نمایش داده می‌شود. در صورتی‌که دبیرخانه اجرایی مسابقه نتایج را بر اساس کل دیتای تست (۱۰۰ درصد) نگهداری کرده و مبنای قضاوت نهایی قرار می‌دهد.

توضیح آنکه: داده‌هایی که به‌منظور تست بر روی سرورها قرار گرفته است، به دو بخش تقسیم می‌شوند. در مرحله میان‌دوره‌ای مسابقه، فقط از ۳۰ درصد داده‌ها به‌منظور تولید دقت کد شما استفاده خواهد شد و در مرحله نهایی، تیم‌هایی انتخاب می‌شوند که روی کل داده‌های تست، دقت بهتری داشته باشد؛ لذا ممکن است تیمی که در مرحله میانی بالاتر از

بقیه تیم‌ها قرار داشته، تیم اول در مرحله نهایی نباشد. اگرچه به احتمال بسیار زیاد، این تیم در بین ۶ تیم برتر خواهد بود.

۱۷- داوری حضوری

در این مرحله، کد نهایی ۶ تیم برتر، مجدداً جمع‌آوری می‌شود. این تیم‌ها در روز اختتامیه iAAA حوزه پزشکی، در حضور کمیته داوری از مدل خود، دفاع خواهند نمود.

شرایط این رویداد، کمیته داوران و... به اطلاع شرکت‌کنندگان رویداد خواهد رسید. در حال حاضر، به‌طور کلی، پارامترهای زیر برای انتخاب تیم‌های برتر جهت دریافت جایزه، اعلام می‌گردد:

- ۷۰ درصد، نمره داوری سیستمی
 - F1 Score
- ۳۰ درصد، نمره داوری حضوری
 - Computation Time
 - Code Review
 - Idea Novelty

۱۸- جدول زمانی

عنوان فعالیت	مهلت انجام
رونمایی از مسئله	۱۳ بهمن ۱۴۰۱
ثبت اطلاعات اعضای تیم	۱ اسفند ۱۴۰۱
انجام تست‌های روان‌شناسی	۱ اسفند ۱۴۰۱
دسترسی به داده‌ها	۴ اسفند ۱۴۰۱
شروع مسابقه	۴ اسفند ۱۴۰۱
ثبت‌نام در پلتفرم لیدر بورد	۶ اسفند ۱۴۰۱
ارسال کدهای دوره‌ای	از ۱۵ اسفند ۱۴۰۱ تا ۵ اردیبهشت ۱۴۰۲
ارسال کد نهایی	۷ اردیبهشت ۱۴۰۲
اعلام نتایج داوری سیستمی	۱۰ اردیبهشت ۱۴۰۲
اختتامیه، داوری و رویداد حضوری	۱۳ اردیبهشت ۱۴۰۲