

طرح مسابقه‌ی دستیار هوشمند صوتی

AI-Bot

از سری مسابقات مساله محور بنیاد ملی نخبگان

AI BOT



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

فهرست مطالب

| | |
|----|--|
| ۲ | ساختار کلی دستیار هوشمند صوتی |
| ۳ | فرآیند کلی مسابقه |
| ۴ | جوایز مسابقه |
| ۴ | منتورینگ |
| ۵ | تیم‌بندی |
| ۵ | توضیحات حوزه های مسابقه |
| ۵ | معرفی بخش سخت افزار |
| ۶ | چالش مقدماتی |
| ۶ | ماموریت چالش مقدماتی |
| ۷ | ورودی‌ها |
| ۷ | ارتباط با سرور |
| ۸ | خروجی‌ها |
| ۸ | قوانین و شرایط چالش مقدماتی |
| ۹ | گزارش طراحی و ساخت دستیار صوتی |
| ۹ | چالش نهایی |
| ۹ | ماموریت تیم‌ها در چالش نهایی |
| ۱۰ | ورودی‌ها |
| ۱۰ | ارتباط با سرور |
| ۱۰ | خروجی‌ها |
| ۱۱ | قوانین و شرایط بخش سخت‌افزار در چالش نهایی |



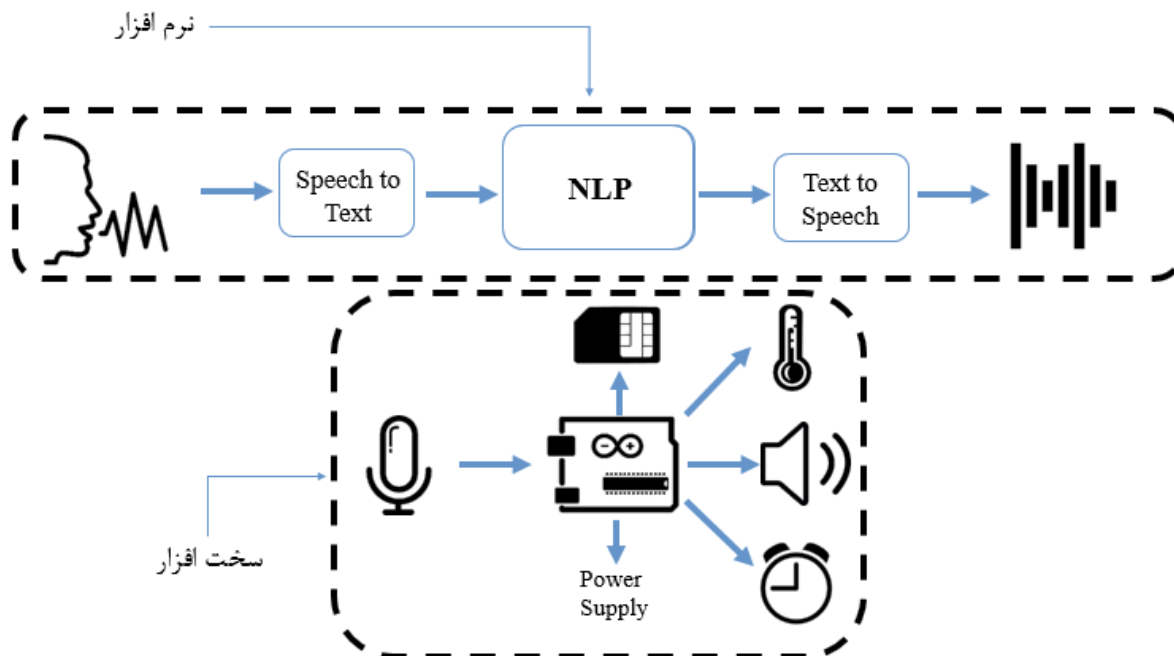
- ۱۱ گزارش طراحی و ساخت دستیار صوتی
- ۱۲ داوری مسابقه
- ۱۲ چالش مقدماتی
- ۱۳ چالش نهایی
- ۱۴ نکات تکمیلی برای بخش سخت افزار
- ۱۵ توضیحات بخش نرم افزار
- ۱۶ چالش مقدماتی نرم افزار
- ۱۶ ماموریت چالش مقدماتی
- ۱۷ انواع فرمان های موجود در چالش مقدماتی
- ۲۲ ورودی و خروجی قسمت نرم افزار در فاز ابتدایی
- ۲۳ فرمت خروجی ها
- ۲۵ سنجه های داوری چالش مقدماتی بخش نرم افزار
- ۲۶ نحوه ی محاسبه امتیاز نهایی در چالش مقدماتی
- ۲۷ چالش نهایی در بخش نرم افزار
- ۲۸ سنجه های داوری چالش نهایی در بخش نرم افزار
- ۲۸ نحوه ی محاسبه امتیاز نهایی در چالش نهایی
- ۲۹ نحوه ی ارسال فایل کدها و استفاده از API نویسا و آریانا
- ۲۹ نحوه ی ارسال فایل کدها روی سرور
- ۳۰ نحوه ی استفاده از API نویسا
- ۳۰ نحوه ی استفاده از API آریانا
- ۳۱ حامیان مسابقه ی AIBOT



دستیار صوتی هوشمند از جمله فناوری‌های به روز دنیاست که کاربردهای بسیاری دارد. از دستیار هوشمند صوتی برای کنترل ابزارهای مختلف مانند تبلت، تلفن، لپ‌تاپ و به طور خاص برای ارسال پیامک، تماس با مخاطب، پخش موسیقی و به طور کلی هوشمندسازی وسایل مختلف استفاده می‌شود. برای این کار کفایت کاربر نیاز خود را از دستگاه بخواهد و به زبان بیاورد، دستیار هوشمند صوتی با گوش کردن به صدای کاربر فرمان موردنظر را انجام می‌دهد. امروزه شرکت‌های داخلی و خارجی بسیاری در زمینه‌ی توسعه این فناوری ورود کرده‌اند که از آن نمونه می‌توان به ¹Alexa، ²Siri، سورنا، دابی اشاره کرد.

ساختار کلی دستیار هوشمند صوتی

به طور کلی دستیار هوشمند صوتی از دو بخش نرم افزار و سخت افزار تشکیل شده است که اجزای مختلف آن در شکل ۱ قابل مشاهده است.



شکل ۱ ساختار کلی دستیار هوشمند صوتی

¹ Amazon

² Apple



همان‌طور که از شکل برمی‌آید در بخش نرم‌افزار، صوت گوینده بایستی طی سه مرحله اصلی به خروجی برسد. در بلوک اول صوت به متن تبدیل شده و سپس به بلوک دوم راه می‌یابد. در بلوک دوم نوع متن تشخیص داده شده و درک می‌شود و پاسخ متناسب با سوال به صورت متنی تولید می‌گردد. پاسخ متنی در مرحله‌ی آخر به بلوک سوم وارد شده و به صوت تبدیل می‌شود. در این مسابقه شرکت کنندگان فقط کافیس بلوک دوم را طراحی نمایند و بلوک اول و سوم یعنی تبدیل صوت به متن و بالعکس به آن‌ها داده می‌شود.

همچنین در قسمت سخت‌افزار، صوت توسط یک میکروفن دریافت شده و به هسته‌ی پردازشی وارد می‌شود. سپس با اتصال به اینترنت پردازش‌های لازم بر روی صوت انجام شده و پاسخ صوتی تولید شده توسط بلندگو پخش می‌شود.

فرآیند کلی مسابقه

مسابقه دارای سه فاز در هر یک از دو حوزه‌ی نرم‌افزار و سخت‌افزار می‌باشد.

فاز اول (آموزش)

در این فاز کارگاه‌های آموزشی مربوط به دو حوزه‌ی سخت‌افزار و نرم‌افزار ارائه خواهد شد. در این کارگاه‌ها، مطالب لازم برای طراحی نرم‌افزار و ساخت سخت‌افزار دستیار صوتی به شرکت کنندگان آموزش داده می‌شود. این آموزش‌ها به صورت آفلاین و از ۲۹ آبان الی ۵ آذر در سایت مسابقه قرار می‌گیرد و استفاده از آن‌ها کاملاً اختیاری است.^۳ لازم به ذکر است شرکت کنندگان در بازه‌ی فوق می‌توانند آموزش‌ها را دریافت نموده و در طول چالش‌ها چندین و چندبار از آن‌ها استفاده نمایند.

فاز دوم (چالش مقدماتی)

فاز دوم چالش مقدماتی است که به صورت مجزا در دو حوزه‌ی طراحی نرم‌افزار و ساخت سخت‌افزار برگزار شده و به مدت یک ماه (۵ آذر الی ۵ دی) طول می‌کشد. موارد خواسته شده در این چالش در ادامه توضیح داده خواهد شد. پس از اتمام این چالش داوری مرحله مقدماتی صورت گرفته و به تیم‌های برتر جوایزی اهدا خواهد شد. لازم به ذکر است به تیم‌های برتر در حوزه‌ی سخت‌افزار در صورت کسب امتیاز لازم، کمک‌هزینه ساخت اعطا می‌گردد.

^۳ بدیهیست که شرکت کنندگان ممکن است برای طراحی پیچیده‌تر و خلاقانه‌تر به آموزش‌های خارج از دروس مذکور نیاز داشته باشند که آموختن آن‌ها به عهده‌ی خود شرکت کننده می‌باشد.



فاز سوم (چالش نهایی مسابقه)

در این فاز کلیه تیم‌های شرکت‌کننده (و نه صرفاً تیم‌های برتر) در فاز دوم مسابقه می‌توانند به صورت ترکیبی از تیم‌های سخت‌افزار و نرم‌افزار تیم‌سازی کرده و در چالش نهایی مسابقه برای تولید یک دستیار هوشمند صوتی کامل وارد شوند. این فاز از ابتدای بهمن شروع شده و تا ابتدای اسفند به طول می‌انجامد. پس از اتمام این چالش داوری نهایی صورت گرفته و به تیم‌های برتر جوایز نقدی و غیر نقدی اهدا خواهد شد.

جوایز مسابقه

باتوجه به چندبخشی بودن مسابقه در فاز مقدماتی برای برگزیدگان بخش نرم‌افزار و سخت‌افزار به صورت جداگانه و در فاز نهایی برای تیم‌های ترکیب شده جوایزی به شرح زیر در نظر گرفته می‌شود.

- جوایز نقدی چالش مقدماتی به تیم‌های اول تا سوم هر حوزه، به ترتیب ۱۰، ۵ و ۲ میلیون تومان
- جوایز نقدی برای چالش نهایی برای سه تیم برتر به ترتیب ۳۰ و ۸ و ۸ میلیون تومان
- اخذ امتیاز نخبگی برای استفاده از تسهیلات بنیاد ملی نخبگان برای چالش نهایی برای سه تیم برتر به ترتیب ۵۰ و ۴۰ و ۳۰ امتیاز (به همه اعضای گروه)
- امکان همکاری تیم‌های برتر با شرکت‌های حامی
- امکان سرمایه‌گذاری بر روی تیم‌های برتر توسط شرکت‌های حامی
- دریافت گواهی‌نامه فنی حرفه‌ای در صورت کسب امتیازهای لازم

منتورینگ

با توجه به زمان کم آموزش و اهمیت برطرف شدن مشکلات آموزشی شرکت‌کنندگان، برای هر حوزه منتوری در نظر گرفته شده است، با توجه به پروتکل تعریف شده در کانال تلگرام، شرکت‌کننده‌ها می‌توانند سوالات خود را از منتور‌ها بپرسند و در جلسات آخر هفته پاسخ آن را دریافت کنند. برای آگاه شدن از زمان و لینک جلسه حتما در کانال تلگرامی عضو شوید.

قابل ذکر است که این جلسات فقط در طول چالش مقدماتی برگزار می‌شود.



تیم‌بندی

در طول مسابقه دو نوع تیم بندی داریم، تیم بندی برای چالش مقدماتی و تیم بندی برای چالش نهایی

تیم‌بندی برای چالش مقدماتی به این صورت است که در هنگام ثبت نام افرادی که می‌خواهند در یکی از دو حوزه ی مسابقه در قالب یک تیم وارد چالش مقدماتی شوند، می‌توانند به صورت تیمی در سایت ثبت نام کنند، به این صورت که سرگروه تیم اسامی و مشخصات اعضای تیم خود را در حین ثبت نام در سایت مسابقه معرفی می کند. با توجه به این تیم بندی در طول چالش مقدماتی با تیم های دیگر در همان حوزه وارد رقابت میشوند.

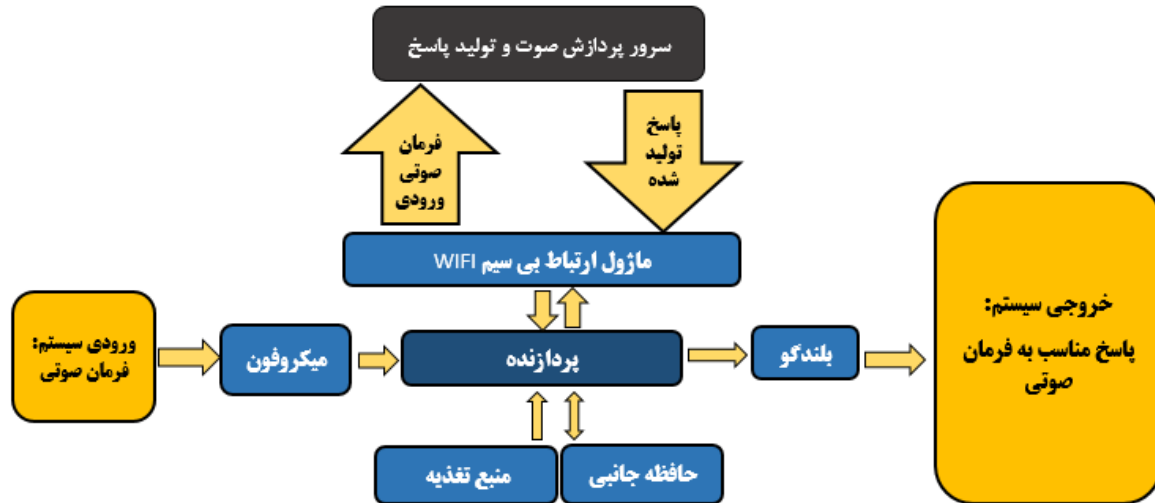
برای تیم بندی بخش چالش نهایی سرگروه های تیم های سخت افزار و نرم افزار باید شناسنامه ای از گروه خود برای معرفی به سایر سرگروه ها آماده کنند، طی پروتکل تعریف شده سرگروه ها باهم آشنا میشوند و میتوانند تیم مناسب تکمیلی خود در حوزه ی دیگر را شناسایی کرده، وارد مذاکره با سرگروه تیم دیگر شوند و در صورت موافقت هر دو سرگروه برای ادغام تیم هایشان، به ادمین تیم سازی اطلاع دهند.

توضیحات حوزه های مسابقه

هر دو بخش نرم‌افزار و سخت‌افزار بایستی طبق خواسته‌های مشخصی طراحی شود. همچنین هر بخش متناسب با پروتکل‌های خاص داوری می‌شوند که در ادامه به توضیح هر بخش می‌پردازیم.

معرفی بخش سخت افزار

در بخش سخت‌افزار مسابقه دستیار صوتی هوشمند، می‌بایست دستگاهی طراحی و ساخته شود که قابلیت دریافت ورودی صوتی، اتصال به اینترنت و ارتباط با سرور، پخش خروجی صوتی و خواندن اطلاعات از روی یک حافظه جانبی را داشته باشد. شکل ۲ نمای کلی قسمت سخت افزار را نمایش می‌دهد.



شکل ۲ نمای کلی بخش سخت افزار

لازم به ذکر است اضافه کردن قابلیت‌های دیگر به موارد مذکور در شکل ۲ ممکن است امتیاز محسوب شود.

چالش‌های مقدماتی

در این مرحله از مسابقه، شرکت‌کنندگان می‌بایست طراحی و ساخت محصول اولیه را انجام دهند. در چالش‌های مقدماتی، مسئله ساده‌سازی شده و لازم است قابلیت‌های کلی و ابتدایی دستگاه پیاده‌سازی شود.

ماموریت چالش‌های مقدماتی

در چالش‌های مقدماتی، تیم‌ها دستگاهی را طراحی و پیاده‌سازی می‌نمایند که یک ورودی صوتی از گوینده دریافت می‌کند و سپس در بستر اینترنت، طبق پروتکل مشخصی با سرور ارتباط گرفته و فایل صوتی ورودی را به سرور ارسال می‌کند. پس از انجام پردازش‌های لازم در سرور، سخت‌افزار فایل تولیدشده را دریافت نموده و صوت خروجی را پخش می‌کند. همچنین دستگاه باید یک حافظه جانبی داشته باشد که تعدادی آهنگ مشخص را روی آن ذخیره نموده و با ورود فرمان آهنگی را پخش نماید. به طور خلاصه، دستیار صوتی در این مرحله باید قابلیت‌های زیر را داشته باشد:

- دریافت ورودی صوتی
- اتصال به اینترنت و ارسال درخواست برای ارتباط با سرور
- ارسال فایل ورودی صوتی به سرور
- دریافت خروجی پردازش‌شده از سرور



- خواندن اطلاعات از حافظه جانبی
- پخش خروجی‌های صوتی

در ادامه توضیحات مربوط به ورودی، خروجی و ارتباط با سرور به طور واضح توضیح داده می‌شوند.

ورودی‌ها

در چالش مقدماتی، ورودی‌های صوتی عبارتند از:

- ✓ "سلام"
- ✓ "آهنگ"

به ازای دریافت هر یک از فرامین فوق و پس از اتصال به سرور و پردازش صوت ورودی، دستوری صادر می‌شود که سیستم بایستی متناسب با آن صوتی را در خروجی منتشر کرده یا آهنگی را از کارت حافظه پخش نماید.

ارتباط با سرور

به دلیل آن که در چالش مقدماتی، تیم‌های سخت‌افزاری با بخش نرم‌افزار مسابقه ارتباط ندارند، به منظور بررسی عملکرد دستگاه‌ها و به دست آوردن نتایج واقعی، یک سرور اولیه توسط برگزارکنندگان مسابقه، راه‌اندازی شده است، که ورودی صوتی به آن ارسال شده، پاسخ‌های متناسب تولید و توسط سخت‌افزار دریافت می‌شود.

برای برقراری ارتباط با این سرور، باید یک درخواست HTTP POST به نشانی

<http://5.160.218.105:8000/AIBotHardware/>

ارسال شود. request body این درخواست باید شامل یک متغیر به نام upload_file باشد که مقدار آن، همان محتوای فایل wav ارسالی (فرمان صوتی ورودی) می‌باشد.

در جواب این درخواست، سرور یک پیام HTTP Response ارسال می‌کند. برای این پیام دریافتی دو حالت وجود دارد.

- اگر فرمان صوتی ارسال شده کلمه آهنگ باشد، content type پیام دریافتی از سرور برابر با text/plain می‌باشد. در این حالت، شماره آهنگی که دستگاه باید پخش کند در قالب یک فایل text ارسال می‌شود.
- content type پیام ارسال شده از سرور، به ازای فرمان سلام، برابر با application/octet-stream می‌باشد. در این حالت دستگاه باید پاسخ صوتی دریافتی را ذخیره کرده و سپس آن را پخش کند.



خروجی‌ها

پس از پردازش صوت در سرور و تولید پاسخ‌های مناسب، دستگاه باید پاسخ را دریافت نموده و عملکرد متناسب با آن را داشته باشد.

- ✓ پاسخ به ورودی "سلام": پاسخ "علیک سلام" در خروجی توسط بلندگو پخش شود.
- ✓ پاسخ به ورودی "آهنگ": در حافظه جانبی، می‌بایست ۱۰ آهنگ که در سایت مسابقه بارگذاری شده است، ذخیره گردد. با دریافت دستور صوتی "آهنگ"، سرور یک شماره تصادفی بین ۱ تا ۱۰ را به سخت‌افزار ارسال می‌کند. سخت‌افزار باید پس از دریافت این عدد، یک آهنگ از میان آهنگ‌های ذخیره‌شده در حافظه را پخش نماید.

قوانین و شرایط چالش مقدماتی

- قابلیت‌های اولیه مذکور بایستی در دستگاه ساخته شده وجود داشته باشد. اضافه کردن هر قابلیت به جز موارد فوق‌الذکر، می‌تواند امتیاز داشته باشد. (نمونه قابلیت‌های اضافی: استفاده از نمایشگر برای نمایش ساعت و تاریخ، ارتباط بلوتوث برای اتصال به تلفن همراه، راه اندازی مازول سیم‌کارت و استفاده از آن در کنار ارتباط WIFI و هر آنچه که به جذابیت دستگاه می‌افزاید).^۴
- در نظر گرفتن منبع تغذیه برای مدار در این مرحله الزامی نیست. تیم‌ها برای تست عملکرد محصول خود می‌توانند از هر روشی برای تغذیه مدار استفاده کنند.
- برای دریافت ورودی صوتی توسط سخت‌افزار، استفاده از بردهایی مانند مازول تشخیص کلام (Voice Recognition Module) در مسابقه ممنوع است.
- جهت اتصال به اینترنت و ارتباط با سرور، به یک مازول ارتباطی نیاز است. در چالش مقدماتی همه تیم‌ها باید از ارتباط WIFI استفاده کنند.
- تمام قطعات مورد استفاده در ساخت سخت‌افزار دستیار صوتی، بایستی قابل تامین از فروشگاه‌های داخل کشور باشد.

^۴ به بخش داوری و امتیازدهی چالش‌ها دقت کنید.



گزارش طراحی و ساخت دستیار صوتی

در انتهای چالش مقدماتی، هر تیم باید یک مستند از چگونگی طراحی دستگاه ارائه دهند. این گزارش می‌بایست شامل موارد زیر باشد:

- ✓ قابلیت های موجود در دستگاه ساخته شده
- ✓ برد پردازنده استفاده شده و دلایل انتخاب آن
- ✓ لیست تمامی قطعات استفاده شده، تعداد هر کدام و قیمت آن ها به دلار

چالش نهایی

پس از داوری چالش مقدماتی، تیم‌ها به فاز بعدی که چالش نهایی است راه می‌یابند. در این مرحله از مسابقه، تیم‌های سخت‌افزار و نرم‌افزار، با یکدیگر ادغام شده و برای ساخت یک محصول کامل که هر دو بخش نرم‌افزار و سخت‌افزار آن کاملاً با هم هماهنگ هستند، تلاش می‌نمایند.

ماموریت تیم‌ها در چالش نهایی

در این بخش از مسابقه، تیم‌ها باید دستگاه خود را ارتقا داده و علاوه بر قابلیت‌های اولیه که در چالش مقدماتی برای دستیار صوتی طراحی و پیاده‌سازی کرده‌اند، موارد دیگری را به آن اضافه کنند. از جمله این که اگر در مرحله قبل اتصالات اجزای مدار با کیفیت مناسب یک محصول واقعی، شکل نگرفته، در این مرحله با چاپ PCB، یک مجموعه با قابلیت اطمینان و پایداری مناسب تولید شود. هم چنین نیاز است که هرگونه خطا از سمت سرور، به کاربر اطلاع داده شود به طوری که در صورت بروز مشکل در ارتباط با سرور، دستگاه اقدامات لازم برای حل مشکل را انجام دهد. مورد دیگری که در چالش نهایی حائز اهمیت است، طراحی و ساخت یک پوشش مناسب برای مدارهای داخلی دستگاه است. این طراحی بایستی توسط نرم‌افزارهای طراحی صنعتی انجام گیرد. (آموزش‌های لازم برای این بخش، قبل از شروع چالش نهایی در اختیار شرکت کنندگان قرار می‌گیرد).

به طور خلاصه، قابلیت‌ها و مشخصات دستگاه نهایی به شرح زیر می‌باشد:

- دریافت ورودی صوتی
- اتصال به اینترنت و ارسال درخواست برای ارتباط با سرور
- ارسال فایل ورودی صوتی به سرور



- دریافت خروجی پردازش شده از سرور
- خواندن اطلاعات از حافظه جانبی
- پخش خروجی های صوتی
- اعلام اشکالات و خطاهای احتمالی در اتصال به سرور
- اتصال قطعات و اجزای مدار روی برد مدار چاپی (PCB)
- مطلوب بودن شکل ظاهری دستگاه (طراحی محصول)

ورودی ها

در این مرحله از چالش، فرامین صوتی ورودی سخت افزار، (برخلاف مرحله پیشین که از سرور دریافت می شدند) از نرم افزار تولید شده توسط تیم نرم افزاری، دریافت می شوند. برای آشنایی با این فرامین صوتی می توانید به بخش مربوطه را در قسمت توضیحات نرم افزار مطالعه کنید. فرامین صوتی در این مرحله به سادگی کلمه "سلام" و "آهنگ" نبوده و جملات کامل در حوزه های گوناگون می باشند.

با دریافت هر فرمان صوتی، سخت افزار باید با سرور ارتباط گرفته و فایل صوتی را به آن ارسال نماید تا پردازش های لازم، در سرور روی آن صورت بگیرد.

ارتباط با سرور

در مرحله مقدماتی، تیم های نرم افزار کدهای خود را در سرور آپلود کرده بودند و یکی از آن کدها را به عنوان کد نهایی انتخاب نمودند. اکنون باید تیم های سخت افزاری به مانند مانند فاز تست سخت افزار در مرحله مقدماتی یک Request Post به آدرس <http://5.160.218.105/AIBotFinal> ارسال کنند. که Files موجود در آن حاوی یک دیکشنری از فایل wav است. با این تفاوت که برای اینکه قرار هست متد تیم نرم افزار فراخوانی شود، کلیدی که باید از دیکشنری فراخوانی شود به صورت username:password است، که username و password به ترتیب یوزرنیم و پسورد تیم نرم افزاری می باشند.

خروجی ها

خروجی فرامین صوتی، متناسب با هر ورودی، با هماهنگی تیم سخت افزار و تیم نرم افزار باید تعیین شود. به عنوان مثال اگر فرمانی مشابه: "یک موسیقی پخش شود" دریافت شد، نرم افزار باید یک فایل قرارداد شده بین دو تیم، به سخت افزار تحویل بدهد تا سخت افزار متوجه شود که باید یک آهنگ از کارت حافظه خوانده و پخش نماید.



قوانین و شرایط بخش سخت‌افزار در چالش نهایی

- قابلیت‌های اولیه دستگاه که در بخش‌های پیشین ذکر شد بایستی در دستگاه ساخته شده وجود داشته باشد. اضافه کردن هر قابلیت به جز موارد فوق‌الذکر، می‌تواند امتیاز داشته باشد. (نمونه قابلیت‌های اضافی: نمایش ساعت و تاریخ، ارتباط بلوتوث برای اتصال به تلفن همراه، راه‌اندازی مازول سیم‌کارت و استفاده از آن در کنار ارتباط WIFI و هر آنچه به جذاب‌شدن دستگاه کمک می‌کند).
- برای دریافت ورودی صوتی توسط سخت‌افزار، استفاده از بردهایی مانند مازول تشخیص کلام (Voice Recognition Module) در مسابقه ممنوع است.
- جهت اتصال به اینترنت و ارتباط با سرور، به یک مازول ارتباطی نیاز است. در چالش نهایی همه تیم‌ها باید از ارتباط WIFI استفاده کنند.
- تمام قطعات مورد استفاده در ساخت سخت‌افزار دستیار صوتی، بایستی قابل تامین از فروشگاه‌های داخل کشور باشند.
- محصول نهایی باید از نظر رابط کاربری و تجربه کاربری (UI و UX) به صورت مطلوبی طراحی شود. به عنوان مثال مجموعه مدارات داخلی، داخل محفظه و پوششی قرار بگیرند که هم از جهت ظاهری زیبایی قابل قبولی داشته و هم از جهت استفاده کاربر از محصول حس خوبی را به او منتقل نماید.
- دستگاه نهایی باید در ابعاد ۷*۷ و با ارتفاع زیر ۱۰ سانتی متر طراحی شود.

گزارش طراحی و ساخت دستیار صوتی

در انتهای چالش نهایی، تیم‌ها باید یک مستند از چگونگی طراحی دستگاه ارائه دهند. این گزارش می‌بایست شامل موارد زیر باشد:

- ✓ قابلیت‌های موجود در دستگاه ساخته شده
- ✓ برد پردازنده استفاده‌شده و دلایل انتخاب آن
- ✓ لیست تمامی قطعات استفاده‌شده و قیمت هر کدام به دلار
- ✓ پیوست فایل‌های طراحی PCB
- ✓ پیوست فایل‌های طراحی محصول (Product Design)
- ✓ پیوست فاکتور خرید قطعات



داوری مسابقه

در انتهای هر دو چالش مقدماتی و نهایی، ارزیابی محصول تیم‌ها انجام می‌شود. در ادامه، نحوه ارزیابی، سنجه‌های داوری و ضرایب ارزیابی برای هر دو مرحله چالش شرح داده شده است.

چالش مقدماتی

در انتهای چالش مقدماتی، تیم‌ها باید محصولی ارائه کنند که قابلیت‌های اولیه ذکر شده در بخش مربوطه را دارد. با توجه به این قابلیت‌ها، سنجه‌ها و ضرایب داوری در این مرحله مطابق جدول ۱ می‌باشد.

| امتیاز | معیارها | |
|--------|--------------|--------|
| ۶۰ | ✓ صحت عملکرد | کارایی |
| ۴۰ | ✓ هزینه ساخت | هزینه |

جدول ۱ ضرایب داوری بخش سخت افزار در چالش مقدماتی

- **صحت عملکرد:** به منظور سنجش عملکرد محصول تیم‌ها، ۲۰ تست (هر تست ۳ امتیاز) روی دستگاه انجام می‌شود. در این ۲۰ تست، دو ورودی سلام و آهنگ با انواع صداها (زن، مرد، کودک، بزرگسال و...)، سطح صدای متفاوت و در فواصل مختلف نسبت به دستگاه، به صورت صوت‌های از پیش ضبط‌شده، پخش می‌گردد. تست در صورتی پاس می‌شود که:
 - پاسخ صحیح متناظر هر ورودی داده شود.
 - در کمتر از ۵ ثانیه پس از اتمام صوت ورودی، صوت خروجی شنیده شود.

- **هزینه ساخت:** به منظور بررسی هزینه ساخت محصول تیم‌ها، شرکت کنندگان موظفند لیست قطعات استفاده‌شده در ساخت محصول را در گزارش خود ارائه دهند. هم‌چنین می‌بایست قیمت قطعات را در واحد دلار ذکر نمایند. کمیته داوری با بررسی قیمت قطعات مورد استفاده، در سایت مرجع (تعیین شده توسط تیم داوری)، صحت هزینه اعلامی توسط تیم‌ها را ارزیابی می‌کنند. در صورتی که خطای تخمین قیمت تیم، بیشتر از ۲۰ درصد قیمت واقعی باشد، آن تیم متخلف شناخته‌شده و از مسابقه حذف می‌گردد. لازم به ذکر است قطعات استفاده‌شده بایستی در بازار قطعات الکترونیکی داخل کشور موجود باشد.



چالش نهایی

در انتهای چالش نهایی، تیم‌ها محصول ارتقا یافته خود را که با بخش نرم‌افزار هماهنگ شده است، ارائه می‌کنند. شرایط داوری این مرحله از دید سخت‌افزاری به شرح زیر است:

| امتیاز | معیارها | |
|--------|--------------------|--------------|
| ۴۰ | ✓ صحت عملکرد | کارایی |
| ۱۵ | ✓ توان مصرفی | |
| ۲۵ | ✓ هزینه ساخت | هزینه |
| ۵ | ✓ پاسخ‌گویی و جذاب | تجربه کاربری |
| ۳ | ✓ کیفیت صدای خروجی | |
| ۵ | ✓ ظاهر فیزیکی | |
| ۲ | ✓ قابل حمل بودن | |
| ۵ | ✓ سایر موارد | |

جدول ۲ ضرایب داوری بخش سخت‌افزار در چالش نهایی

▪ **صحت عملکرد:** به منظور سنجش عملکرد محصول تیم‌ها، ۲۰ تست (هر تست ۲ امتیاز) روی دستگاه انجام می‌شود. در این ۲۰ تست، ورودی‌های تعریف‌شده در بخش نرم‌افزار، با انواع صداها (زن، مرد، کودک، بزرگسال و...)، سطح صدای متفاوت و در فواصل مختلف نسبت به دستگاه، به صورت صوت‌های از پیش ضبط شده، پخش می‌گردد. تست در صورتی پاس می‌شود که:

- پاسخ صحیح متناظر هر ورودی داده شود.
- در کمتر از ۵ ثانیه پس از اتمام صوت ورودی، صوت خروجی شنیده شود.

▪ **توان مصرفی:** به منظور سنجش دستگاه از نظر توان مصرفی، لازم است دستگاه ساخته‌شده، یک سوکت جهت اتصال منبع تغذیه خارجی داشته باشد. کمیته داوری با اندازه‌گیری جریان مصرفی، در حین انجام تست شماره ۱، توان مصرفی دستگاه را اندازه‌گیری می‌نماید.

▪ **هزینه ساخت:** به منظور بررسی هزینه ساخت محصول تیم‌ها، شرکت‌کنندگان موظفند لیست قطعات استفاده‌شده در ساخت محصول را در گزارش خود ارائه دهند. همچنین می‌بایست قیمت قطعات را در واحد دلار ذکر نمایند. کمیته داوری با بررسی قیمت قطعات مورد استفاده، در سایت مرجع (تعیین شده توسط



کمیته داوری، صحت هزینه اعلامی توسط تیم‌ها را ارزیابی می‌کنند. در صورتی که خطای تخمین قیمت تیم، بیشتر از ۲۰ درصد قیمت واقعی باشد، آن تیم متخلف شناخته شده و از مسابقه حذف می‌گردد. لازم به ذکر است قطعات استفاده‌شده بایستی در بازار قطعات الکترونیکی داخل کشور موجود باشند.

- **پاسخ گویا و جذاب:** این قسمت در همکاری با بخش نرم‌افزاری تیم انجام می‌شود.
- **کیفیت صدای خروجی:** صدای خروجی باید برای شنوندگان مطلوب و تا حد ممکن بدون نویز و صدای ناخوشایند باشد. همچنین سطح صدای خروجی نیز باید مناسب باشد و به آسانی توسط کاربر شنیده شود. میانگین امتیاز کمیته داوری، امتیاز این بخش را تشکیل می‌دهد.
- **ظاهر فیزیکی:** ارزیابی این معیار با نظر اعضای کمیته داوری انجام می‌شود و میانگین امتیاز ایشان به عنوان امتیاز این معیار در نظر گرفته می‌شود.
- **قابل حمل بودن:** در صورتی که برای دستگاه یک باتری مناسب تعبیه شود، امتیاز این بخش به تیم تعلق می‌گیرد. البته هزینه باتری استفاده‌شده در لیست هزینه تیم‌ها لحاظ نمی‌شود.
- **سایر موارد:** در نظر گرفتن قابلیت‌هایی که تجربه کاربری مطلوبی برای کمیته داوری (به عنوان کاربر محصول) ایجاد نماید، با نظر این کمیته، امتیاز خواهد داشت.

نکات تکمیلی برای بخش سخت‌افزار

احتمال می‌رود در طراحی و پیاده‌سازی دستیار صوتی هوشمند، شرکت‌کنندگان با چالش‌هایی مواجه شوند. تعدادی از چالش‌های احتمالی که لازم است افراد در فرآیند طراحی محصول خود، در نظر بگیرند، در این بخش شرح داده شده است.

✓ دریافت ورودی‌های صوتی با نویز بالا: در صورتی که کیفیت صوت خوانده‌شده توسط دستگاه مناسب نباشد و سیگنال، نویز زیادی داشته باشد، ممکن است سرور پردازش صوت با مشکل مواجه شود و نتواند صوت ورودی را تشخیص دهد. (چالش مقدماتی و نهایی)



- ✓ نرخ ارسال صوت به سرور: فایل صوتی ارسال شده به سرور تبدیل صوت به متن (نویسا) باید به فرمت wav و با نرخ ۱۶ ۱۶ bit KHz ارسال شود. ممکن است برای دریافت و ارسال صوت با این نرخ، محدودیت سخت افزاری در بعضی پردازنده ها و یا ماژول ها وجود داشته باشد. (چالش مقدماتی و نهایی)
- ✓ تشخیص ورودی معتبر: ممکن است دستگاه با هر صدایی در محیط، آماده به پاسخ شود. برای رفع این مشکل می توان (به عنوان مثال) یک اسم مشخص برای دستگاه تعریف کرد که با ورود آن، دستگاه آماده پاسخ دهی شود البته از روش های جایگزین ابتکاری هم می توانید استفاده کنید. (چالش نهایی)
- ✓ ورود فرمان صوتی در زمان پخش آهنگ: ممکن است قبل از اتمام آهنگ، یک فرمان صوتی وارد شود. (چالش نهایی)

لازم به ذکر است که موارد فوق، تمام نکات و چالش های محتمل را در بر ن گرفته و ممکن است موارد دیگری نیز وجود داشته باشد که بایستی در طراحی و پیاده سازی دستگاه (در هر دو بخش سخت افزار و نرم افزار) مورد توجه قرار گیرند.

توضیحات بخش نرم افزار

دستیارهای هوشمند با گرفتن جمله ای در قالب زبان طبیعی به عنوان ورودی، اطلاعاتی را استخراج کرده، سپس پردازشی انجام می دهند و پاسخ پرسش دریافتی را تولید می کنند. به طور کلی اطلاعاتی که باید استخراج شوند را می توان در ۲ دسته معرفی کرد:

(۱) تشخیص فرمان (Intent Classification): پس از دریافت جمله، نوع فرمان را باید تعیین کنید.

(۲) استخراج آرگومان ها (Slot filling): آرگومان های مربوط به جمله و سوال مطرح شده را کنید.

مثال: جمله ی زیر را در نظر بگیرید:

«فردا هنگام اذان ظهر هوای تهران چگونه است؟»

دستیار هوشمند پس از مشاهده ی جمله ی فوق ابتدا باید تشخیص دهد، سوال مربوط به فرمان داده شده در مورد آب و هوا است. سپس باید آرگومان های این فرمان را استخراج کند. آرگومان های این سوال موارد زیر هستند:

تاریخ - وقت شرعی - تهران (به عنوان شهر)



حال که فرمان به طور کامل تشخیص داده شده است و آرگومان‌های مورد نظر سوال استخراج شده‌اند، دستیار هوشمند می‌تواند تابع یا API مربوط به فرمان دریافت شده را فراخوانی کند. کیفیت یک دستیار هوشمند در ۳ مورد بررسی خواهد شد:

✓ تشخیص دادن نوع فرمان

✓ استخراج آرگومان‌های مرتبط

✓ تولید پاسخ مدنظر فرمان صوتی.

دقت دستیار هوشمند در هر یک از این ۳ مورد به عنوان معیار سنجش در نظر گرفته می‌شود.

در این بخش ابتدا به جزئیات و طراحی مسابقه و معیارهای داوری خواهیم پرداخت.

چالش مقدماتی نرم افزار

به طور خلاصه چالش ابتدایی را می‌توان به صورت زیر تعریف کرد:

“دریافت تعدادی فرمان متنی به عنوان ورودی و استخراج جزئیات متناسب با هر فرمان متنی.”

لازم به ذکر است که چالش حائز اهمیت در این مرحله مربوط به پردازش متن است. در نتیجه بازشناسی و سنتز گفتار، توسط برنامه‌ای آماده در اختیار شرکت کنندگان قرار می‌گیرد. داوری نهایی روی مجموعه دادگان به صورت آفلاین انجام می‌شود. همچنین شرکت کنندگان برای انجام این مرحله ۱ ماه زمان خواهند داشت.

ماموریت چالش مقدماتی

در مرحله‌ی ابتدایی آنچه مدنظر است، ساخت یک عامل هوشمند است که بتواند با دریافت فرمان‌هایی که تحت بستر زبان طبیعی بیان می‌شوند، بتواند فرمان‌ها را تشخیص داده و آرگومان‌های مرتبط را استخراج کند.

فرمان‌هایی که به عامل هوشمند داده می‌شوند تک‌نوبتی هستند و عامل هوشمند باید در یک‌نوبت تمامی اطلاعات مربوط به فرمان را بیان کند، به بیان دیگر عامل هوشمند برای فهمیدن منظور پرسشگر نمی‌تواند سوالی بپرسد.

کاری که عامل هوشمند باید انجام دهد، دریافت فرمان و برگرداندن موارد زیر می‌باشد:

(۱) تشخیص نوع فرمان

(۲) استخراج آرگومان‌ها



۳) پاسخ به سوال مطرح شده

۴) آدرس API فراخوانی شده

انواع فرمان‌های موجود در چالش مقدماتی

در جدول زیر می‌توان لیست تمامی فرمان‌های مربوط به چالش مقدماتی را به همراه آرگومان‌های مربوط به هر فرمان مشاهده کرد.

| شماره فرمان | عنوان فرمان | آرگومان‌ها | | | |
|-------------|-------------|-------------|-------------|-------------|---------------|
| | | آرگومان اول | آرگومان دوم | آرگومان سوم | آرگومان چهارم |
| ۱ | آب و هوا | شهر | تاریخ | ساعت | - |
| ۲ | اوقات شرعی | نوع اوقات | شهر | تاریخ | - |
| ۳ | ساعت | شهر | - | - | - |
| ۴ | تقویم | نوع تقویم | تاریخ | مناسبت | - |

جدول ۳ لیست فرمات‌های مربوط به چالش مقدماتی

در ادامه هر کدام از این فرمان‌ها را با ذکر مثال توضیح خواهیم داد. در نظر داشته باشید که این مثال‌ها صرفاً برای فهم بهتر فرمان‌های مورد نظر هستند. لذا سوالات مربوط به هر فرمان می‌تواند پیچیده‌تر از مثال‌های مطرح شده باشد.

فرمان صوتی ۱: آب و هوا

این فرمان مربوط به آب و هوا بوده و آرگومان‌هایی که از سوال مطرح شده باید استخراج شوند، نام شهر، تاریخ و ساعت (در صورت وجود) هستند.

| جمله | آرگومان اول | آرگومان دوم | آرگومان سوم |
|---|-------------|-------------|-------------|
| آب و هوای فردای تهران چطور است؟ | تهران | ۶ آذر | - |
| فردا ساعت ۶ بعد از ظهر هوای شیراز چگونه است؟ | شیراز | ۶ آذر | ۱۸:۰۰ |
| دمای هوای مسکو در روز جمعه ۹ آبان چند درجه است؟ | مسکو | ۹ آبان | - |

جدول ۴ نمونه سوالات و آرگومان‌ها در رابطه با فرمان صوتی ۱



سوالات مربوط به آب و هوا از API مربوط به سایت <https://openweathermap.org> طراحی خواهند شد. لذا لازم است که مستندات مربوط به این API را مطالعه کنید و واژگان مربوط به نام شهرها را فارسی سازی کنید. آنچه که شما می توانید به کمک این API به صورت رایگان به دست آورید می تواند مورد سوال قرار بگیرد، که این موارد شامل دمای شهر (برحسب درجه سلسیوس)، سوال از داده های تاریخی، سوال از پیش بینی های آینده و ... است.

مثلا از لینک

<http://api.openweathermap.org/data/2.5/forecast?q=tehran&appid=31f22d9862c83c9cb2aa02687bb25c87>

می توانید مستندات مربوط به استخراج پیش بینی دمای هوای مربوط به ساعات آینده را ۳ ساعت به ۳ ساعت مطالعه کنید. یک نمونه استفاده از این API مربوط به پیش بینی دمای هوای تهران را می توانید از لینک زیر مشاهده کنید :

<http://api.openweathermap.org/data/2.5/forecast?q=tehran&appid=31f22d9862c83c9cb2aa02687bb25c87>

* تذکر ۱ : زمانی که سوالی مرتبط به آب و هوا پرسیده می شود (داده ی تاریخی یا پیش بینی) منظور نزدیک ترین داده ی موجود نسبت به زمان مطرح شدن پرسش است. مثلا اگر سوال در ساعت ۱۸:۳۰ پرسیده شود و داده ای برای این ساعت ثبت نشده باشد، به گزارش های آب و هوا مربوط به ساعات دیگر نگاه می شود. اگر در ساعات ۱۲:۰۰ و ۱۹:۴۰ گزارشی ثبت شده باشد، گزارش مربوط به ساعت ۱۹:۴۰ انتخاب خواهد شد، چون به زمان پرسیده شدن سوال نزدیک تر است.

* تذکر ۲ : در صورتی که در متن سوال، به صراحت ساعتی برای پیش بینی آب و هوا و یا داده ی تاریخی ارائه نشد، ساعت ۱۲ ظهر را به عنوان ساعت پیش فرض برای استخراج دما در نظر بگیرید.

فرمان صوتی ۲: اوقات شرعی

جدول زیر را در نظر بگیرید :

| جمله | آرگومان اول | آرگومان دوم | آرگومان سوم |
|---|-----------------|---------------|-------------|
| اذان ظهر تهران چه موقعی است؟ | اذان ظهر | تهران | ۶ آذر |
| نیمه شب شرعی تورنتو چه زمانی است؟ | نیمه شب شرعی | تورنتو | ۶ آذر |
| اذان مغرب فردای قم چه ساعتی است؟ | اذان مغرب | قم | ۶ آذر |
| فاصله ی اذان مغرب تهران و اذان صبح مشهد | اذان مغرب تهران | اذان صبح مشهد | - |

جدول ۵ نمونه سوالات و آرگومان ها در رابطه با فرمان صوتی ۲



سوالات مربوط به اوقات شرعی، از API سایت aladhan.com طرح می‌شوند. لذا مستندات مربوط به آن باید مطالعه شوند، همچنین در پیاده‌سازی شرکت کنندگان واژگان خاص مربوط به این API به فارسی برگردانده شوند. ممکن است از تمامی خروجی‌های ممکن این API سوال طرح شود، لذا شرکت کنندگان در پیاده‌سازی‌های خود باید آمادگی تعامل با هرگونه خروجی از این API را داشته باشند. پارامترهای مربوط به فراخوانی از این API محدود به شهر، کشور، روز، ماه و سال می‌باشد و پارامترهای دیگر (مانند روش محاسبه اوقات شرعی (method) و مکتب فقهی (school)) مورد سوال قرار نمی‌گیرند.

<https://aladhan.com/prayer-times-api#GetCalendarByCitys>

برای فراخوانی این API نیاز به وارد کردن اسم کشورها، نام شهرها و روز و ماه و سال نیز هست. در نتیجه شرکت کنندگان باید روی استخراج نام شهرها و کشورها نیز فعالیت کنند و برای استخراج نام شهرها و کشورها می‌توانند از هر منبعی که می‌خواهند استفاده کنند. به مثال زیر توجه کنید :

<http://api.aladhan.com/v1/hijriCalendarByCity?city=tehran&country=ir&method=2&month=4&year=2020&day=7>

در این مثال اوقات شرعی تهران در تاریخ ۲۰۲۰-۰۴-۰۷ پرسیده شده است و پاسخ دریافت شده شامل اوقات شرعی ۹ گانه است. در نتیجه فراخوانی این API برای پیدا کردن یک وقت شرعی خاص مانند اذان ظهر نیز می‌تواند کارایی داشته باشد. برای جواب سوال از این API نیازی به تطبیق محدوده زمانی Time Zone به اوقات محل مورد پرسش نیست.

فرمان صوتی ۳: ساعت

سوالات مربوط به این فرامین را می‌توان به کمک API موجود در <http://worldtimeapi.org> پاسخ داد. به جدول زیر که حاوی سوالات مربوط به این دسته است توجه کنید.



| جمله | آرگومان اول | آرگومان دوم |
|--|-------------|-------------|
| ساعت چند است؟ | تهران | - |
| الان در نیویورک ساعت چند است؟ | نیویورک | - |
| در مونیخ چه زمانی از روز است؟ | مونیخ | - |
| اختلاف ساعت تهران و مسکو چند ساعت است؟ | تهران | مسکو |

جدول ۶ نمونه سوالات و آرگومان‌ها در رابطه با فرمان صوتی ۳

فرمان صوتی ۴: تاریخ

| جمله | آرگومان اول | آرگومان دوم |
|---|-----------------|-------------|
| تاریخ شمسی امروز چیست؟ | شمسی | - |
| امروز به میلادی چه روزی است؟ | میلادی | - |
| تاریخ ۲۰۲۰-۰۶-۰۵ به قمری چه تاریخی است؟ | قمری | ۲۰۲۰-۰۵-۰۶ |
| چند روز تا روز جهانی دیابت مانده است؟ | روز جهانی دیابت | - |
| دوشنبه هفته‌ی بعد به تاریخی هجری؟ | دوشنبه هفته بعد | هجری |
| مناسبت ۲۷ رجب چیست؟ | ۲۷ رجب | - |

جدول ۷ نمونه سوالات و آرگومان‌ها در رابطه با فرمان صوتی ۴

برای استخراج داده‌ی مربوط به تقویم لازم است شرکت کنندگان API مربوط به تقویم را خودشان توسعه دهند. منبع این دادگان وبسایت time.ir است. در بخش مستندات که شامل کد منبع نیز می‌شود حتما بررسی می‌شود که جواب‌های مربوط به سوالات حتما حاصل crawl و استخراج اطلاعات از time.ir باشد و نباید از API و یا وبسایت و یا فایل دیگری استخراج شود.

نیازی به پیاده‌سازی API دادگان crawl شده نیست. دادگان خزش شده را می‌توان به هر طریقی داخل کد منبع استفاده کرد.

در خزش وبسایت time.ir باید موارد زیر استخراج شود و امکان سوال از این موارد وجود دارد: روز، ماه، سال، به سه تاریخ میلادی، شمسی و قمری به همراه مناسب‌های این ایام در هر سه تقویم میلادی، شمسی، هجری و تعطیلات رسمی مناسبی.



وبسایت time.ir برای تعیین مناسبت‌های جهانی (میلادی) و مذهبی (قمری) و روزهای ملی (شمسی) از فرمت مشخصی تبعیت می‌کند.

crawl لحظه‌ای وبسایت برای پاسخ به سوالات ضرورتی ندارد بلکه بهتر است حاصل crawl از وبسایت در یک دیتابیس محلی (Local) یا فایل ذخیره شود. سوالات مربوط به تقویم از ۵ سال اخیر و ۵ سال آینده پرسیده می‌شود.

****تذکر ۱:** همچنین امکان طرح سوالات پرسش و پاسخ منطقی نیز وجود دارد در پاسخ به این پرسش‌ها مستلزم اعمال عملگرهای مقایسه، حداقل، حداکثر، اختلاف، میانگین از بین دو یا چندین اطلاع خواسته شده می‌باشد.

مثال: از بین شیراز و اهواز و قزوین کدامیک فردا ساعت ۱۸ سردتر است؟

مثال: اختلاف دمای بین سردترین و گرمترین ساعت روز در مشهد کدام است؟

مثال: فاصله‌ی اذان مغرب امروز با نماز صبح فردا؟

****تذکر ۲:** ممکن است پاسخ به یک سوال نیازمند فراخوانی از دو API باشد. مثلاً «در اولین تعطیل رسمی پیش‌رو دمای هوا در ساعت ۱۶:۳۵ چند درجه است؟» یافتن جواب این سوال نیازمند دو فراخوانی است: فراخوانی از API تقویم برای پیدا کردن اولین تعطیلی رسمی و سپس فراخوانی از API آب و هواست. توجه شود در چنین مثالی سوال فقط یک intent دارد و آن هم آب و هواست، با اینکه پاسخ این سوال مستلزم فراخوانی دو API است ولی intent همان موضوع مورد سوال است. این دست سوالات از دسته‌ی سوالات پرسش و پاسخ ترکیبی هستند که دو موضوع با هم ترکیب می‌شوند و نسبت به سوالاتی که درخواست مستقیم از یک API دارند از درجه سختی بالاتری برخوردارند.

****تذکر ۳:** همچنین سوالاتی ممکن است بطور همزمان هم ترکیبی و هم منطقی باشد که پیچیده‌ترین و سخت‌ترین درجه سختی سوالات است.

مثال: دمای هوا در ساعت ۱۶:۳۰، روز جهانی دیابت تهران کمتر است یا عید مبعث؟

مثال: سردترین اذان ظهر هفت روز گذشته در شیراز چه روزی بوده است؟

توجه: توصیه می‌شود شرکت کنندگان در وقت محدود یک ماهه تمرکز خود را ابتدا بر سوالات عادی سپس منطقی و ترکیبی و در نهایت سوالات منطقی-ترکیبی قرار دهند که به ترتیب درجه سختی سوالات در آنها افزایش می‌یابد.



طرح سوالات از دسته‌های پیچیده تر برای جداسازی تیم‌های برتر است، لذا به منطق طرح سوالات چالش به شکل مرحله به مرحله بنگرید.

****تذکر ۴:** دسته‌بندی intent سوالات باید بصورت open-world باشد، در صورتیکه intent سوال طرح شده جز چهار دسته‌ی آب و هوا، تقویم، ساعت، اوقات شرعی نیست باید دسته سوال طرح شده را به unknown برچسب بزنند. مثلاً اگر از قیمت ارز سوال شود چون intent این سوال در این چهار مورد نمی‌گنجد باید unknown خروجی داده شود.

****تذکر ۵:** معادل‌سازی و برگردان اسم‌های خاص مانند نام شهرها و کشورها از فارسی به انگلیسی برای فراخوانی API ها بر عهده تیم‌هاست که می‌توانند از هر منبعی به اختیار خود استفاده کنند.

****تذکر ۶:** واحد پردازشگر متن باید طوری طراحی شود که آرگومان‌های ورودی را در هر فرمتی شناسایی و درک کند. بطور مثال ساعت ۴ بعد از ظهر می‌تواند با ۱۶:۰۰ و یا چهار غروب تعبیر شود.

****تذکر ۷:** در صورتی که آرگومان‌های سوال برای پاسخ دادن به آن ناکافی است باید جوابی مبنی بر ناکافی بودن اطلاعات سوال برگرداند.

****تذکر ۸:** امکان وجود چند آرگومان از یک نوع (مثلاً دو آرگومان از اوقات شرعی) در یک جمله وجود دارد. در نتیجه اگر چند آرگومان در صورت سوال وجود داشت، برنامه باید تمامی آنها را بازگرداند. در بخش فرمت خروجی‌ها می‌توانید مشاهده کنید که در صورت وقوع چنین حالتی چگونه باید آرگومان‌های کشف شده را ثبت کنید.

ورودی و خروجی قسمت نرم افزار در فاز ابتدایی

ورودی به دستیار هوشمند یک جمله در قالب زبان طبیعی است. خروجی دستیار هوشمند نیاز بایستی یک json باشد که شماره فرمان و آرگومان‌های آن در آن آمده باشد. برای مثال به مثال‌های زیر دقت کنید:

ورودی: آب و هوای فردای تهران ساعت ۱۸:۳۵ چه طور است؟



خروجی مطلوب:

```
{
  'city'      :      ["تهران"],
  'time'      :      ['18:35'],
  'date'      :      ['فردا'],
  'type'      :      ['1'],
  'result'    :      '29'
  , 'api_url':      '...'
}
```

ورودی: اذان ظهر به افق ساری چه وقت است؟

خروجی مطلوب:

```
{
  'religious_time' :      ["اذان ظهر"],
  'city'          :      ['ساری'],
  'type'          :      '2',
  'result'        :      '13:08'
  , 'api_url':      '...'
}
```

فرمت خروجی‌ها

- پس از تشخیص نوع فرمان، باید عددی را به صورت string بازگردانید. کلید این فرمان “type” خواهد بود. عدد مربوط به هر فرمان به صورت زیر است:

(۱) آب و هوا (۲) اوقات شرعی (۳) ساعت (۴) تقویم

یعنی اگر نوع فرمان از ساعت بود، عدد ۳ را برمی گردانیم.

**** تذکر:** در صورتی که سوال پرسیده شده مربوط به هیچ کدام از فرمان‌های فوق نبود، در دسته‌ی unknown قرار می‌گیرد. لذا در این حالت “1-” برگردانید.

- برای بازگرداندن آرگومان‌ها، کلیدها را به صورت زیر نامگذاری کنید.

city ❖
date ❖
time ❖
religious_time ❖
calendar_type ❖
event ❖



■ هر کدام از این کلیدها حاوی یک لیست هستند که شامل آرگومان‌های مربوط به این کلیدها است. مقادیر مربوط به هر کلید توسط یکی از موارد زیر پر می‌شود. اگر کلیدی وجود داشت که آرگومانی برای آن موجود نبود، مقدار مربوط به آن کلید را با یک لیست خالی پر کنید.

✓ آرگومان‌های مربوط به اوقات شرعی باید در کلید religious_time قرار بگیرند، مقدار این آرگومان‌ها به صورت زیر است، یعنی در آرگومان‌هایی که باز می‌گردانیم عیناً یک یا چند مورد از موارد زیر را باید برگردانیم:

امساک، اذان صبح، طلوع آفتاب، اذان ظهر، اذان عصر، غروب آفتاب، اذان مغرب، اذان عشا، نیمه شب شرعی

✓ برای آرگومان‌های مربوط به نام شهر از کلید city استفاده می‌کنیم. نام شهر به صورت یک رشته متشکل از حروف فارسی بدون کاراکتر اضافی است، مانند “تهران”. نام شهرهای موجود را به لیست موجود در کلید city اضافه می‌کنیم.

✓ برای آرگومان‌های مربوط به ساعت از کلید time استفاده می‌کنیم. ساعت‌ها باید با فرمت hh:mm:ss ساعت برگردانده شوند. برای مثال ۱۳:۲۰:۰۰ یعنی ساعت ۱ و ۲۰ دقیقه.

✓ آرگومان‌های مربوط به تاریخ در کلید date قرار می‌گیرند. این آرگومان‌ها باید از فرمت yyyy-mm-dd تبعیت کنند. مثال: ۱۳۹۹-۰۸-۲۹

✓ نوع تقویم در کلید calendar_type نگهداری می‌شود. مقادیر مربوط به آن می‌تواند “میلادی”، “قمری” و یا “شمسی” باشد.

✓ نام مناسبت در کلید event نگهداری می‌شود. نام مناسبت موجود در تقویم باید به صورت یک رشته به لیست مربوط به این کلید اضافه شود، مانند “عید مبعث”.

■ کلیدی با نام api_url تعریف کنید، در صورتی که از API استفاده می‌کنید، فقط و فقط لینک یک API را به همراه پارامترهای آن را توسط یک رشته به این کلید اختصاص دهید. در غیر اینصورت مقدار این کلید خالی تعریف می‌شود.

■ در نهایت کلید result حاوی پاسخی است که دستیار هوشمند شما، پس از پردازش سوال ورودی به عنوان جواب باز می‌گرداند. پاسخ باید به یکی از صورت‌های زیر باشد:

❖ **تاریخ:** یک رشته با فرمت yyyy-mm-dd بازگردانده شود. مثال: ۱۳۹۹-۰۸-۲۹

❖ **ساعت:** یک رشته با فرمت hh:mm:ss بازگردانده شود. مثال: ۲۲:۳۰:۰۰

❖ **عدد:** یک رشته حاوی عدد مورد نظر بازگردانده شود.



❖ **مناسبت** : یک رشته حاوی نام مناسبت (یا مناسبت‌ها) بازگردانده شود. در صورتی که چند مناسبت بازگردانده می‌شود آنها را به کمک کاراکتر - از یکدیگر جدا کنید.

❖ **آب و هوا** : در صورتی که قرار است نوع آب و هوا مشخص شود، نام مربوط به وضعیت آب و هوا برگردانده شود. نوع آب و هوا را به یکی از ۵ صورت زیر برگردانید :
ابری - بارانی - آفتابی - برفی - طوفانی

** تذکر : در صورتی که با اطلاعات داده شده در صورت سوال، نمی‌توان پاسخی تولید کرد، مقدار مربوط به کلید result را برابر یک string خالی قرار دهید.

سنجه‌های داوری چالش مقدماتی بخش نرم‌افزار

عملکرد هر یک از دستیارهای هوشمند روی دادگان تست سنجیده می‌شود و معیار برتری دستیار هوشمند میزان دقت آن در تشخیص فرمان و آرگومان‌های آن می‌باشد. در واقع امتیاز نهایی هر دستیار هوشمند، ترکیبی از دقت آن روی سه پارامتر تشخیص فرمان، پرکردن آرگومان‌های فرمان، پاسخ فرمان است. به طور کلی سنجه‌های ارزیابی تیم‌ها در گام اول عبارتند از:

✓ دقت ماژول پردازش متن در تشخیص فرمان‌ها

به تشخیص‌های صحیح امتیاز داده می‌شود. در صورتیکه تشخیص فرمان صحیح نباشد و برای بخش‌های بعدی یعنی پرکردن آرگومان‌ها و صحبت خروجی جواب‌ها ارزیابی صورت نمی‌گیرد نمره بخش‌های دیگر صفر خواهد بود.

✓ پرکردن آرگومان‌ها (ارزیابی کمی)

فرمول محاسبه امتیاز این بخش:

تعداد آرگومان درست تشخیص داده شده = Correct

تعداد آرگومان تشخیص داده نشده = Incorrect

تعداد آرگومان تشخیص داده شده اضافی = Additional Arguments

Score = Max(0, Correct-Incorrect-Additional Arguments)

مثال: فرض کنیم آرگومان‌های یک سوال عبارت از A، B و C باشند. در صورتی که تشخیص A و B باشد امتیاز حاصل به صورت زیر محاسبه می‌شود:

$$\text{Score} = \text{Max}(0, 2-1-0) = \text{Max}(0, 1) = 1$$

و در صورتیکه که تشخیص A، B، D و E باشد، ۲ آرگومان صحیح و ۱ آرگومان غلط و ۱ عدد آرگومان اضافه داریم. در نتیجه امتیاز حاصل به صورت زیر محاسبه می‌شود:

$$\text{Score} = \text{Max}(0, 2-1-1) = 0$$



توجه: در صورتیکه امتیاز منفی شود، امتیاز صفر در نظر گرفته می شود.

✓ صحت خروجی جوابها

فرمت های خروجی پرسش ها باید با فرمت استاندارد اعلام شده برای کلید result برابری کند. فقط تطبیق عینی پاسخ شرکت کننده با پاسخ صحیح مشمول صحت می شود .
**تذکر ۱: محصول این فاز بایستی به گونه ای آماده شود که بتواند هر چه بیشتر فرمان های مربوطه را تشخیص دهد به این معنا که برای یک فرمان می توان از عبارت بندی های مختلفی استفاده کرد که قاعدتا همه این عبارت بندی ها در مثال های ذکر شده بالا نیامده است.

**تذکر ۲: استفاده از API نویسا اختیاری است. (نیاز به توضیح کامل تر)

نحوه محاسبه امتیاز نهایی در چالش مقدماتی

تعداد n سوال از برنامه ی شرکت کنندگان پرسیده خواهد شد. هر سوال شامل ۱۰۰ امتیاز است :

| نام معیار | توضیح | امتیاز (از ۱۰۰) |
|-----------|------------------------|-----------------|
| Intent | تشخیص صحیح intent سوال | ۳۰ |
| Arguments | تشخیص صحیح آرگومان ها | ۳۰ |
| Result | تشخیص صحیح پاسخ نهایی | ۴۰ |

جدول ۸ امتیاز بندی چالش مقدماتی

در صورت تشخیص اشتباه intent امتیاز ۰ در نظر گرفته می شود. به صورت فرمولی امتیاز هر سوال مانند سوال i (Q_i) را می توان به صورت زیر در نظر گرفت :

$$Score(Q_i) = \begin{cases} 0 & \text{if } intent_i = 0 \\ intent_i + arguments_i + result_i & \text{otherwise} \end{cases}$$

تعداد سوالات در چالش نهایی n سوال است. در نتیجه میانگین امتیازات تمامی سوالات محاسبه می شود :

$$Final_score = \sum_{i=1}^n \frac{Score(Q_i)}{n}$$

که این عدد، عددی بین ۰ تا ۱۰۰ خواهد بود.



چالش نهایی در بخش نرم افزار

پس از به دست آوردن مازول پردازش زبان در گام قبلی، در فاز دوم بایستی این مازول به مازولهای تبدیل صدا به متن و تبدیل متن به صدا متصل شوند. به این صورت که سیستم کلی حاصل، صدایی را از ورودی گرفته آن را تحویل به مازول تبدیل صدا به متن نموده و متن ورودی را به دست می آورد. سپس، این متن به مازول پردازش زبان (که در فازمقدماتی طراحی شد) داده می شود و این مازول نوع فرمان و آرگومانهای آن را تشخیص می دهد و در نهایت پاسخ مناسب تولید کند. برای مثال در صورتی که نوع فرمان اوقات شرعی و آرگومانهای آن نماز صبح، تهران، ۳ دی باشد؛ این مازول بایستی در خروجی خود با استفاده از فراخوانی های api های موجود برای هر فرمان، تولید کند که نماز صبح تهران در ۳ دی لحظه ۰۷:۴۰:۰۵ می باشد. سپس این متن خروجی به مازول تبدیل متن به صدا داده شده و خروجی صوتی سیستم تولید می شود.

چالش های فاز نهایی شامل موارد زیر می باشد.

- مدیریت کردن چالش خطای مازول تبدیل صوت به متن و انتشار خطا به فاز پردازش زبان ممکن است ورودی مازول متن شامل انواع خطا لفظی، املائی، جمله بندی و ... باشد. مازول متن باید این خطا را تحمل کند و با وجود خطا در جمله نوع فرمان، آرگومان ها و جواب جمله را به صحت برگرداند.
 - طراحی و پیاده سازی مازول تولید متن
 - به جهت ارزیابی توسعه پذیری راه حل تیم های برگزیده در مرحله ی مقدماتی در ۱۰ روز پایانی مرحله ی نهایی شیوه نامه در اختیار شرکت کنندگان قرار می گیرد که از تیم ها خواسته می شود حوزه ی جدید مانند چهار حوزه ی مقدماتی به فرمان های دستیار صوتی اضافه کنند.
- **تذکر:** احتمال تغییر در موارد چالش نهایی و اضافه شدن چالش های دیگر نیز وجود دارد. در صورت تغییر قبل از چالش نهایی اطلاعی رسانی خواهد شد.
- **تذکر:** متصل کردن مازول های بازشناسی و سنتز گفتار به صورت یک برنامه آماده به شرکت کنندگان داده خواهد شد و از این رو شرکت کنندگان درگیر راه اندازی این دو مازول نخواهند بود. البته لازم به ذکر است که شرکت کنندگان به دلخواه می توانند از هر مازول بازشناسی و سنتز دیگری که علاقه مند هستند نیز به جای مازول های ارائه شده استفاده کنند.



سنجه‌های داوری چالش نهایی در بخش نرم‌افزار

داوری این بخش هم بصورت انسانی و هم به صورت ماشینی می‌باشد و در بخشی از عملکرد تیم‌ها بسته به نظرات داوران در رابطه با عملکرد صحیح سیستم سنجیده می‌شود و در باقی عملکردها که نظام ارزیابی کمی برای آن ممکن است امتیاز دهی کمی مشابه مرحله‌ی مقدماتی خواهد بود. از جمله سنجه‌هایی که در این بخش مورد توجه قرار می‌گیرند می‌توان به موارد زیر اشاره کرد:

- میزان جذابیت پاسخ‌های تولیدشده توسط مازول تولید متن بطور مثال جواب‌هایی که در قالب جمله کامل هستند جذابیت بیشتری نسبت به جواب‌های کوتاه دارند. بررسی امتیازدهی به خروجی تولید شده توسط تیمی از داوران و بصورت subjective انجام خواهد شد.
- کیفیت مستندات و گزارش تولیدشده توسط تیم برای حوزه‌ی جدید که در قبل از چالش نهایی به سامانه دستیار صوتی اضافه می‌شود یک ارزیابی کمی دقیقاً مانند آنچه در مرحله‌ی مقدماتی بیان شده صورت خواهد گرفت.
- پرسش‌های از هر پنج حوزه (چهار حوزه‌ی مرحله‌ی مقدماتی + حوزه‌ی مرحله‌ی جدید مرحله‌ی نهایی) به همراه انواع خطا لفظی، املائی و جمله‌بندی به مازول متن داده می‌شود و ارزیابی‌های مرحله‌ی مقدماتی شامل تشخیص فرمان، تشخیص آرگومان‌ها و تطبیق جواب با همان نظام امتیازبندی اعمال می‌شود.

نحوه‌ی محاسبه امتیاز نهایی در چالش نهایی

تعداد n سوال از برنامه‌ی شرکت‌کنندگان پرسیده خواهد شد، نظام امتیازدهی برای تشخیص نوع فرمان، تشخیص آرگومان‌ها و تطبیق جواب، از نظام امتیازدهی چالش ابتدایی تبعیت می‌کند، امتیاز این بخش ۷۰ از ۱۰۰ است. ۳ معیار داوری دیگر نیز در این بخش اضافه شده است، نحوه‌ی امتیازدهی به هر بخش را می‌توانیم در جدول زیر ببینیم.



| نام معیار | توضیح | امتیاز (از ۱۰۰) |
|-----------------|-----------------------------------|-----------------|
| intent | تشخیص صحیح intent سوال | ۲۰ |
| arguments | تشخیص صحیح آرگومان‌ها | ۲۰ |
| result | تشخیص صحیح پاسخ نهایی | ۳۰ |
| speed | سرعت اجرا | ۱۰ |
| interestingness | میزان جذابیت پاسخ‌های تولید شده | ۱۰ |
| reporting | کیفیت مستندات و گزارشات تولید شده | ۱۰ |

جدول ۹ امتیازبندی داوری چالش نهایی

معیارهای جذابیت پاسخ‌های تولید شده، کیفیت مستندات و گزارشات تولید شده و سرعت اجرا به صورت subjective و توسط داوران بررسی می‌شوند. امتیاز نهایی در این بخش نیز از ۱۰۰ خواهد بود.

نحوه‌ی ارسال فایل کدها و استفاده از API نویسا و آریانا

برنامه‌ی شما در فاز نهایی باید با قسمت سخت‌افزاری تعامل داشته باشد، که این باعث این می‌شود که نیاز داشته باشیم از API تبدیل گفتار به متن و API مربوط به تبدیل متن به گفتار استفاده کنیم. در این مسابقه امکان استفاده از API های نویسا و آریانا فراهم شده است. اگر تیمی قصد استفاده از API های دیگری را دارد، می‌تواند از آنها استفاده کند. در نتیجه با دقت موارد زیر را مطالعه کنید.

نحوه‌ی ارسال فایل کدها روی سرور

ابتدا فایل‌های خود را که با پسوند py هستند، درون یک فایل zip قرار دهید. فایلی که سرور آن را اجرا خواهد کرد را AIBOT.py بنامید. به ۲ صورت می‌توانید برنامه‌ی خود را بنویسید:

(۱) در صورتی که قصد استفاده از آریانا و نویسا را به صورت همزمان دارید، درون فایل AIBOT.py یک تابع با نام AIBOT تعریف کنید که ورودی آن یک string است که حاوی متن سوال است و آنچه که باز می‌گرداند یک json است که چگونگی ساخت آن را در قسمت فرمت خروجی‌ها تعریف کرده‌ایم. در این حالت نیازی به پیاده‌سازی قسمت خاصی برای ارتباط با نویسا و آریانا ندارید، چون این امکان درون سرور برای شما به صورت پیش‌فرض تعریف شده است.

(۲) در صورتی که قصد استفاده از آریانا و یا نویسا را ندارید، درون فایل AIBOT.py یک تابع با نام AIBOT_Modified تعریف کنید که ورودی آن یک string است که حاوی آدرس فایل صوتی مربوط به سوال



است و خروجی آن یک json است که چگونگی ساخت آن را در قسمت فرمت خروجی ها تعریف کرده ایم. در این حالت قسمت تبدیل صوت به متن و یا قسمت تبدیل متن به گفتار را باید خودتان پیاده سازی کنید و یا از API مدنظر خود استفاده کنید. اگر قصد استفاده از نویسا و یا آریانا را در قسمتی از برنامه ی خود داشتید، می توانید به صورت زیر از آنها استفاده کنید :

نحوه ی استفاده از API نویسا

فایلی با نام nevisa.py در اختیار شرکت کنندگان قرار می گیرد. در این فایل یک تابع به نام nevisa وجود دارد. این تابع ۳ پارامتر ورودی دارد.

(۱) فایل صوتی با فرمت wav

(۲) یک string شامل ip

(۳) یک comment که باید به string منحصر به فرد باشد. (ترجیحا در این بخش از کد ملی خود استفاده کنید)
فایل را به صورت زیر بخوانید، سپس تابع nevisa را به صورت زیر فراخوانی کنید :

```
file = open("name_of_file.wav", "rb")  
text = nevisa(file, ip, comment)
```

متغیر text حاوی پرسش موجود در صدای فرستاده شده خواهد بود.

نحوه ی استفاده از API آریانا

فایلی با نام ariana.py در اختیار شرکت کنندگان قرار می گیرد. در این فایل یک تابع به نام ariana وجود دارد. این تابع ۲ پارامتر ورودی دارد.

(۱) متنی که قرار است تبدیل به گفتار شود.

(۲) آدرس فایلی که قرار است صوت تولید شده در آن نگهداری شود.

پس از فراهم کردن ۲ پارامتر فوق، تابع ariana را به صورت زیر فراخوانی کنید :

```
ariana('your text', 'path/to/output.wav')
```

پس از انجام این کار فایل صوتی مورد نظر تولید می شود و در آدرسی که مشخص شده است، نگهداری می شود.



حامیان مسابقه‌ی AIBOT

- ✓ بنیاد ملی نخبگان
- ✓ شرکت فناوری
- ✓ شرکت فولاد هیربد
- ✓ شرکت همراه اول
- ✓ ستاد اقتصاد دیجیتال و هوشمند سازی
- ✓ ماد (مرکز توسعه مهارت های اقتصاد دیجیتال)
- ✓ شرکت عصر گویش پرداز
- ✓ مسابقه ی ره نشان
- ✓ ستاد فرهنگ سازی اقتصاد دانش بنیان (معاونت علمی و فناوری ریاست جمهوری)