ششمین کنفرانس بین‌المللی
«بازی‌های رایانه‌ای؛ فرصت‌ها و چالش‌ها»

DeepSkill : یک چارچوب جهت رتبه‌بندی مهارت تیم‌ها در بازی‌های برخط چندنفره انبوه

محمد مهدی رضاپور[1*]، محمدعلی نعمت‌بخش[2]، افسانه فاطمی[3]

۱- دانشجوی دکتری نرم‌افزار دانشگاه اصفهان
Mru.x90@gmail.com

۲- استاد تمام گروه مهندسی نرم‌افزار دانشگاه اصفهان
nematbakhsh@eng.ui.ac.ir

۳- استادیار گروه مهندسی نرم‌افزار دانشگاه اصفهان
a_fatemi@eng.ui.ac.ir

چکیده

در سال‌های اخیر با همه‌گیری بسترهای مبتنی بر شبکه‌های جمعیتی انبوه نظیر شبکه‌های اجتماعی و پیام‌رسان‌ها بسیاری از خدمات رایانه‌ای دیگر نیز با تاثیرپذیری از این پارادایم تحولاتی را از مناظر مختلف تجربه کرده‌اند. از جمله موفق‌ترین بسترهای مبتنی بر این پارادایم را می‌توان بازی‌های برخط چند نفره انبوه (بَچنا) دانست. یکی از مسائلی که هدف بسیاری از پژوهش‌های حوزه‌ی بازی‌های دیجیتال قرار گرفته است چگونگی جذب هر چه بیشتر بازیکنان انسانی به بازی است. در واقع سوال اصلی مطرح شده این است که چگونه می‌توان در حین اجرای بازی، بازیکنان انسانی را به اصطلاح در دنیای بازی غرق نمود. یکی از اصلی‌ترین مقولات تاثیرگذار بر رضایت بازیکنان در بازی‌های بچنا قرار گرفتن در مقابل بازیکنانی با مهارت مشابه است که در این صورت احتمال برد و باخت او متناسب باشند. در حال حاضر روش‌هایی چون EloRating، TrueSkill و TrueSkill2 به منظور سنجش و رده‌بندی مهارتی بازیکنان در بازی‌های چند نفره برخط ارائه شده است که با توجه به تعریف بازی‌های بچنا برای استفاده در این نوع از بازی‌ها پیش‌بینی‌های لازم را نداشته‌اند؛ از همین رو در این پژوهش تلاش می‌شود یک چارچوب با عنوان DeepSkill جهت سنجش و رتبه‌بندی مهارتی بازیکنان در بازی‌های بچنا ارائه شود. طبق نتایج به دست آمده DeepSkill توانسته با رویکردی نوین به تعریف مهارت در بازی‌های بَچنا خطای بسیار کمی در تخمین مهارت تیمی در اینگونه از بازی‌ها از خود نشان دهد.

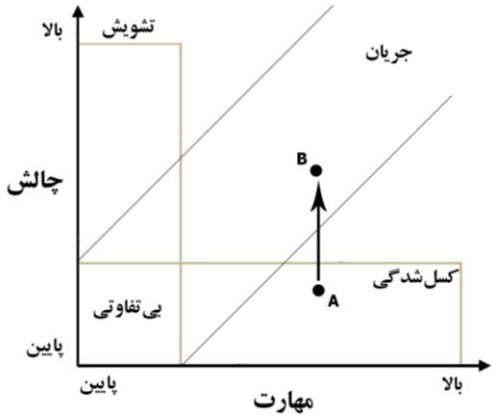
کلمات کلیدی: سنجش مهارت، بازی‌های MMO، رده‌بندی مهارتی، سیستم‌های چندعامله، DeepSkill

۱–مقدمه

در حال حاضر یکی از دغدغه‌های اصلی مجامع پژوهشی و صنعتی در حوزه‌ی بازی‌های رایانه‌ای جذب حداکثری بازیکنان به بازی و اطمینان از وفاداری آن‌ها است[۱]. تاثیر مطلوب و مورد نظر در رابطه با جذابیت بالای بازی‌ها در ادبیات مسئله با عنوان غرق شدگی نیز مطرح می‌گردد[۲]. میهای چیکسِنت‌میهای [۳] روانشناس نامدار چنین حالت روانی را اولین بار با عنوان وضعیت

جریان[۱]معرفی کرد. او در [۴] کیفیت تجربه‌ی بازیکن انسانی در بازی را تابعی از دو متغیر میزان چالش‌های بازی[۲]و میزان مهارت بازیکن[۳]معرفی کرده است.

چیکسنت‌میهای و الیس [۵] دسته‌بندی و چگونگی توزیع کیفیت تجربه‌ی بازی را در قالب دو مدل چهار کاناله و هشت کاناله ارائه داده‌اند که در حال حاضر مدل چهارکاناله مبنای اصلی بسیاری از پژوش‌های حوزه بازی‌ها قرار دارد و این مدل در شکل ۱نشان داده شده است. این مدل حالات احساسی انسان‌ها را در چهار وضعیت بی تفاوتی، جریان، تشویق، و کسل شدگی دسته بندی کرده است. طبق این مدل در صورتی که بازیکن در حین بازی در نقطه A قرار داشته باشد نشان می دهد که او با میزان چالش و سختی کمی به نسبت با مهارت ها و توانایی هایش مواجه شده است و لذا نیاز است که بازی با توجه به میزان مهارت بالای او چالش بیشتری را پیش روی او قرار دهد(یا به عبارت دیگر سخت تر شود) تا وضعیت این بازیکن از نقطه A به نقطه B تغییر یابد.



شکل ۱ مدل چهار کاناله مبتنی بر نظریه جریان

با توجه به نقش بسیار مهم سنجش هر چه دقیق تر مهارت بازیکنان انسانی برای مشخص شدن جایگاه بازیکن در مدل چهار کاناله جریان و پیچیدگی‌های بسیار زیاد چنین سنجشی، این پژوهش با نگاهی نو ابعاد مسئله‌ی قرار دادن بازیکن در کانال جریان را به سنجش مهارت بازیکن در بازی کاهش داده است.

بازی‌های برخط چندنفره انبوه[۴](بچنا) که با دو شاخصه‌ی منحصر به فرد مانایی جهان بازی و اجتماعی بودن بافت بازی به عنوان یکی از پرمخاطب‌ترین نوع بازی‌های رایانه‌ای مطرح شده و به زعم طیف وسیعی از پژوهشگران همپوشانی بسیار زیادی میان مهارت‌های به کار گرفته شده در این بازی‌ها با مهارت‌های قرن ۲۱ وجود دارد[۶] [۷]، نیازمند چارچوب نوینی در مبحث سنجش مهارت بازیکنان است.

راهکارهایی چون TrueSkill [۸]و TrueSkill2 [۹] به عنوان مطرح‌ترین سیستم‌های سنجش و رتبه‌بندی بازیکنان در بازی‌های چند نفره برخط[۵]که برای استفاده در پلتفرم Xbox Live مایکروسافت ارائه شده اند پیشبینی‌های لازم جهت پوشش طیف

[۱] Flow state
[۲] Challenges
[۳] Skills
[۴] Massively multiplayer online games
[۵] Multiplayer online games

وسیع مهارتی در بازی‌های بچنا را نداشته‌اند. همچنین الگوریتم α-Rank[۱۰] با رویکردی مبتنی بر نظریه بازی‌ها با چالش‌های چون عدم مقیاس‌پذیری به ازای تعداد بالای بازیکنان و استراتژی‌های آن‌ها و همچنین عدم پوشش طیف مهارتی بازی‌های بچنا در نسخه فعلی روبرو است.

یکی از مشخصه‌های بازی‌های بچنا را می‌توان فعالیت‌های تیمی دانست و بسیاری از بازیکنان نیز تمایل به پیش‌بردن بازی به صورت تیمی دارند. لذا در این پژوهش تلاش شده است که به جای مبنا قرار داده شدن مهارت فردی بازیکنان، مهارت تیمی در نظر گرفته شود. نظر به اهمیت وجود راهکاری برای سنجش مهارت تیم‌ها در بازی‌های بچنا، این پژوهش در نظر دارد یک چارچوب برای چنین سنجشی ارائه نماید. در همین راستا تلاش می‌شود به سه سوال زیر پاسخ داده شود.

۱. چگونه می توان مهارت تیم‌ها در بازی های بچنا را در فضای چند بعدی مدل نمود؟
۲. در صورتی که این مدل سازی انجام شود چگونه می توان سنجش مهارت تیم‌ها را هنگامی که چندین روش خبرگی وجود دارد انجام داد؟
۳. چگونه می‌توان مهارت تیمی بازیکنان را حتی در صورت به اتمام نرساندن بازی اندازه‌گیری نمود؟

در ادامه این مقاله و در بخش دوم پیشینیه‌ی کارهای انجام گرفته بررسی خواهد شد، در بخش سوم روش پیشنهادی این پژوهش با عنوان DeepSkill ارائه شده و در بخش چهارم نتیجه‌گیری از مطالب ارائه خواهد شد.

۲-کارهای پیشین

سنجش مهارت و رتبه بندی بازیکنان در بازی‌های رایانه‌ای دارای ادبیات غنی در بخش آموزش و پراکنده در حوزه‌های دیگر است. از جمله پژوهش هایی که در زمینه بازی های آموزشی انجام گرفته است می توان به [۱۲, ۱۳, ۱۴, ۱۵] اشاره نمود. شوت در [۱۶] اولین بار نشان داد که چگونه می توان با استفاده از رویکرد طراحی سنجش مبتنی بر شواهد[۱]سنجش مهارت را به گونه ای انجام داد که بازیکن به عنوان یادگیرنده در محیط یک بازی آموزشی متوجه سنجیده شدن نشود و آن را سنجش نهان[۲]نامید. مطابق با این روش مسئله ی سنجش به سه مدل با عناوین مدل عمل، مدل شواهد و مدل توانایی تقسیم می گردد. مدل عمل شامل تمامی اعمالی که بازیکن در بازی انجام می دهد می شود. مدل توانایی شامل تمامی متغیرهای ذهنی که قرار است سنجیده شود شده و در نهایت مدل شواهد می بایست مقادیر متغیرهای مدل وظیفه را به نحوی به مقادیری برای متغیرهای مدل توانایی تبدیل نماید. او روش خود را در یک بازی آموزشی با نام Taiga مورد ارزیابی قرار داد. در این پژوهش به منظور تشکیل مدل شواهد، تنها به استفاده از ابزاری با نام JMap برای نمره دهی به جواب های دانش آموزان اکتفا شده است. شوت همچنین در کارهایی چون [۱۷, ۱۸] با روشی مشابه توانایی های دانش آموزان در بازی را مورد سنجش قرار داده است.

یکی از مهمترین کارهای مورد نیاز جهت سنجش مهارت بازیکنان در بازی های بچنا، مشخص شدن مجموعه متغیرهای توانایی مورد نظر جهت سنجش است. در واقع با تعبیری که در متدولوژی ECD با عنوان مدل توانایی شناخته می شود، هر سنجش موفقی نیازمند شناسایی صحیح و کامل متغیرهای مهارتی نهایی جهت سنجش است. سورملیز [۷] با مطالعه ۱۲۰ مقاله در زمینه مهارت های موثر در بازی های بچنا به این نتیجه رسیده است که بازی های بچنا توانایی پرورش عمده مهارت های قرن ۲۱ را دارا هستند. سورملیز به نقل از بینکلی [۱۹] که با مطالعه مروری جامع خود در زمینه مهارت های قرن ۲۱، به

[۱] Evidence centered assessment design
[۲] Stealth assessment

یک چارچوب واحد با عنوان KSAVE رسیده است، این چارچوب را به عنوان مبنای مهارت‌های موثر در بازی های بچنا معرفی می کند. ابعاد چهارگانه و مهارت های ده گانه چارچوب KSAVE به شرح زیر است:

بعد ۱: مهارت‌های فکری
- خلاقیت و نوآوری
- تفکر منتقدانه، توانایی حل مسئله، قدرت تصمیم گیری
- یادگیری برای یادگیری، فراشناخت

بعد ۲: شیوه انجام کار
- ارتباطات
- همکاری(کار تیمی)

بعد ۳: کار کردن با ابزار
- سواد اطلاعاتی
- سواد فناوری اطلاعات

بعد ۴: زندگی در جهان
- شهروندی–محلی و جهانی
- زندگی و حرفه
- مسئولیت های شخصی و اجتماعی

مکرری در [۲۰] با توسعه نظریه وِب [۲۱] که چهار حوزه مرتبط با عملکرد ذهنی در بازی های بچنا را معرفی کرده است، یک مدل ۵ بعدی شامل موقعیت مکانی(برای مثال جست و جوی محیط)، تخصیص مکانی[۱](برای مثال نمای نمایش بازی)، تعامل مکانی[۲](برای مثال تعامل با المان های هوش مصنوعی)، ادراک مکانی[۳](برای مثال فهم قوانین، امکانات در دسترس و خصوصیات محیطی) و تعامل مکانی–اجتماعی[۴](برای مثال استفاده از ابزارهایی چون چت گروهی)، ارائه می کند. مکرری سپس به منظور سنجش مهارت‌های بازیکنان در بازی دنیای وارکرفت که یک بازی بچنا محسوب می شود، هر کدام از این ابعاد را با مشخص سازی تعداد سنجه و در قالب یک ماتریس رفتاری جهت انجام محاسبات بعدی کمی سازی می نماید. برای مثال سنجه‌هایی چون استفاده از نکات[۶]در بازی، استفاده از ابزارهای مختلف بازی، تقاضا برای کمک و نهایتا استفاده از سیستم راهنمای بازی به منظور کمی سازی بعد "ادراک مکانی" استفاده شده‌اند.

راهکارهایی چون TrueSkill که در سال ۲۰۰۷ توسط هربریچ و همکاران از پژوهشگران مایکروسافت ارئه شده است[۸]، بازیکنان را بر مبنای یک الگوریتم بیزینی و با توجه به مهارت آن‌ها رده‌بندی می‌کنند. نتیجه این رده‌بندی به منظور رویارویی بازیکنان برابر از نظر مهارتی در یک اجرا از بازی استفاده می‌شود. اگرچه ایده مطرح شده در این روش توسط مایکروسافت پتنت انحصاری شده است، اما به علت کاربرد بسیار گسترده این روش در پلتفرم‌های بازی مطرح این شرکت چون Xbox live، در مجامع پژوهشی بسیار مورد توجه بوده است. یکی از محدودیت‌های TrueSkill بسنده کردن به تاریخچه برد و باخت

[۱] Spatial Positioning
[۲] Spatial Appropriation
[۳] Spatial Interactivity
[۴] Spatial Realization
[۵] Socio-Spatial Interactivity
[۶] Hints

بازیکنان برای تخمین مهارت آن‌هاست لذا تمهیدی برای در نظر گرفتن مهارت بازیکنانی که به دلایل غیر قابل پیشبینی قادر به پایان رساندن بازی نبوده‌اند نداشته است. همچنین دقت این روش در حدود ۷۰ درصد بوده است و از طرف دیگر به صورت مستقیم نمی‌توان از آن به عنوان روشی جهت تخمین مهارت بازیکنان استفاده نمود.

نسخه دوم TrueSkill که در سال ۲۰۱۸ توسط مایکروسافت ارائه شده است [۹] تا حدی سعی کرده است ضعف TrueSkill در عدم پوشش ویژگی‌های عملکردی بازیکنان در بازی را با در نظرگیری ویژگی‌هایی چون امتیازتجربه‌بازیکن، عضویت در گروه‌ها، تعداد کشته‌ها، تمایل به خروج از بازی که در بازی‌های سبک تیراندازی یافت می‌شوند، محاسبه دقیق‌تری از مهارت بازیکنان در قیاس با نسخه اول TrueSkill انجام بدهد. اما با وجود این پیشرفت می‌توان گفت همچنان TrueSkill2 با محدودیت جدی در قبال بازی‌های غیر سبک تیراندازی مواجه است و نمی توان از آن به عنوان یک راهکار مناسب در دیگر سبک‌های بازی بهره گرفت. البته می توان گفت در سبک تیر اندازی هم ویژگی های عملکردی متعددی را می توان برشمرد که در TrueSkill2 دیده نشده اند. از دیگر ایرادات وارد به این روش می‌توان به شیوه‌ی اثر دادن ویژگی‌های عملکردی بازیکنان اشاره نمود؛ به این صورت که صرفا با در نظر گیری قوانینی ایستا و ذکر شروطی در رابطه با مقادیر این ویژگی‌ها، تغییراتی در ارزیابی‌های عملکردی اعمال می‌شود.

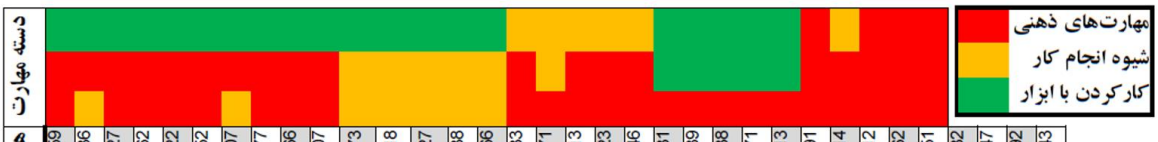
۳-راهکار پیشنهادی

بر اساس آن‌چه که در بخش‌های پیشین ارائه شد، در این پژوهش تلاش می‌شود که در قالب چارچوبی با عنوان DeepSkill، به سه سوال اصلی مطرح شده این پژوهش پاسخ داده شود. فرایند کلی سنجش در این چارچوب بر مبنای متدلوژی طراحی سنجش مبتنی بر شواهد (ECD) شکل گرفته است که در لایه "مدل توانایی" آن از چارچوب KSAVE استفاده خواهد شد. دلیل این انتخاب سازگاری این چارچوب با تعریف مهارت‌های موثر در بازی های بچنا بوده است. همچنین نظر به دقت و مقیاس‌پذیری بسیار مطلوبی که روش‌های مبتنی بر شبکه‌های عصبی عمیق در سال‌های اخیر از خود نشان داده‌اند، ما نیز در این پژوهش از ظرفیت یادگیری عمیق در لایه شواهد استفاده خواهیم نمود.

به منظور ارزیابی عملکرد DeepSkill از مجموعه داده های یک بازی چند نفره برخط با عنوان[۱] PUBG استفاده شده است. این بازی که از محبوبیت بسیار بالایی در میان کاربران بازی های چند نفره برخط برخوردار است(رده سوم بیشترین تعداد بازیکن در میان بازی‌های برخط با ۶۰۰ میلیون بازیکن) از نظر برخی کارشناسان یک بازی بچنا محسوب نمی شود، اما بسیاری از خصوصیات این بازی‌ها را به خصوص در بخش تعاملات تیمی دارد. هر مسابقه از این بازی در یک محیط مشترک و برای ۱۰۰ بازیکن روایت می‌شود. این بازیکنان می‌توانند در تیم‌های تک یا چند عضوی قرار گرفته و بازی را شروع کنند. هدف در این بازی زنده ماندن است. در نهایت رتبه بندی تیم‌ها در این بازی بر اساس ماندگاری آن‌ها در زمین بازی مشخص می‌شود. بر این اساس طی تحلیل‌های صورت گرفته و مشاوره نخبگانی، نگاشتی میان چارچوب DeepSkill و ویژگی‌های مجموعه داده نهایی انجام شده است. این نگاشت شامل سه بعد اصلی از چارچوب DeepSkill که عبارت‌اند از "مهارت‌های ذهنی"، "شیوه انجام کار" و "مهارت کار با ابزار" شده است(جدول۱). در ادامه مراحل پیکربندی چارچوب DeepSkill جهت انجام ارزیابی‌های عملکردی تشریح خواهد شد.

[۱] PlayerUnknown's Battlegrounds

جدول۱ لیست متغیرهای پیش‌بین در کنار میزان همبستگی آن‌ها با متغیر هدف و همچنین نمایش دسته‌بندی آن‌ها بر مبنای مدل توانایی چارچوب DeepSkill



شرح متغیر	متغیرهای پیشین	همبستگی با رتبه	دسته مهارت
مجموع تعداد بازیکن کشته	player_kills_sum	-0.442997559	
میانگین تعداد بازیکن کشته	player_kills_mean	-0.475717836	
حداقل تعداد بازیکن کشته	player_kills_min	-0.3665027	
حداکثر تعداد بازیکن کشته	player_kills_max	-0.454639362	
انحراف معیارتعداد بازیکن کشته	player_kills_std	-0.395627022	
مجموع میزان صدمه زده	player_dmg_sum	-0.441558552	
میانگین میزان صدمه زده	player_dmg_mean	-0.502242007	
حداقل میزان صدمه زده	player_dmg_min	-0.402525877	
حداکثر میزان صدمه زده	player_dmg_max	-0.466037066	
انحراف معیار میزان صدمه زده	player_dmg_std	-0.388534807	
مجموع کمک رسانی به همتیمی	player_assists_sum	-0.278574273	
میانگین کمک رسانی به همتیمی	player_assists_mean	-0.294999518	
حداقل کمک رسانی به همتیمی	player_assists_min	-0.180431727	
حداکثر کمک رسانی به همتیمی	player_assists_max	-0.283696138	
انحراف معیار کمک رسانی به همتیمی	player_assists_std	-0.342217766	
مجموع قرار گرفتن در حالت نیمه جان	player_dbno_sum	-0.250425733	
میانگین قرار گرفتن در حالت نیمه جان	player_dbno_mean	-0.263721871	
حداقل مقدار قرار گرفتن در حالت نیمه جان	player_dbno_min	-0.160029913	
حداکثر مقدار قرار گرفتن در حالت نیمه جان	player_dbno_max	-0.239442323	
انحراف معیار قرار گرفتن در حالت نیمه جان	player_dbno_std	-0.318411846	
مجموع فاصله رانندگی	player_dist_ride_sum	-0.4586331	
میانگین فاصله رانندگی	player_dist_ride_mean	-0.530387239	
حداقل فاصله رانندگی	player_dist_ride_min	-0.486477538	
حداکثر فاصله رانندگی	player_dist_ride_max	-0.512779771	
انحراف معیار فاصله رانندگی	player_dist_ride_std	-0.308782713	
مجموع فاصله پیادروی	player_dist_walk_sum	-0.326772791	
میانگین فاصله پیادروی	player_dist_walk_mean	-0.327826614	
حداقل فاصله پیادروی	player_dist_walk_min	-0.354005112	
حداکثر فاصله پیادروی	player_dist_walk_max	-0.209358962	
انحراف معیار فاصله پیادروی	player_dist_walk_std	-0.061017651	
تعداد تیم‌ها در بازی	game_size	-0.022571132	
تعداد همتیمی‌ها	party_size	0.021712147	
تاریخ انجام بازی	date_norm	0.003938392	
زمان انجام بازی	time_norm	-0.0014771043	

۱-۳ آماده سازی مجموعه داده PUBG

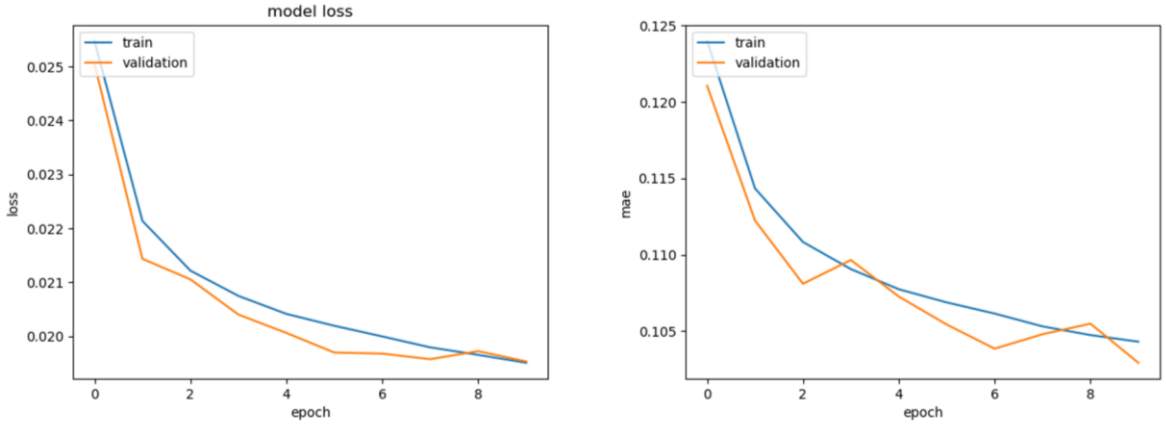
مجموعه داده اولیه PUBG معادل ۱۸ گیگابایت داده حاصل از ۷۲۰ هزار مسابقه است که با روش‌های سمپلینگ استاندارد مجموعه داده‌ای شامل اطلاعات ۱۱ هزار مسابقه استخراج شد. مجموعه سمپل شده شامل ۱ میلیون تاپل داده‌ای با ۱۶ متغیر است که هر تاپل شامل ویژگی‌های عملکردی یک بازیکن در یک تیم مشخص و در یک مسابقه مشخص است. از میان ۱۶ متغیر اولیه، متغیر رتبه تیمی در پایان هر مسابقه، به عنوان متغیر پیش‌بین انتخاب شد. به منظور شناخت بیشتر از ویژگی‌ها، الگوها و ارتباطات موجود در این مجموعه داده، از متد[1]EDA استفاده شد. نهایتا با انجام مجموعه‌ای از عملیات‌های تحلیلی و تجمعی که منجر به استخراج ۴۲ متغیر از ۱۶ متغیر اولیه شد، مجموعه داده عملکردی تیمی فراهم شد. در مجموعه داده تیمی استخراج شده تلاش شده است که تا جای ممکن تمامی ابعاد عملکرد تیمی دیده شود(جدول۱). پس از آن نرمال‌سازی مجموعه متغیرها به روش Min-Max نیز انجام پذیرفت. همچنین متغیر هدف(رتبه‌تیمی) با در نظر گرفتن متغیر "تعداد تیم‌ها" و در نهایت روش Min-Max نرمال‌سازی شد. به منظور مستقل ساختن سنجش مهارت تیم‌ها از میزان زمان بازی‌کرد بازیکنان، مقادیر متغیرهای عملکردی فردی هر کدام از بازیکنان بر اساس میزان زمان حضور آن‌ها در بازی نرمال‌سازی شد و لذا متغیر زمان در مجموعه داده نهایی پیش‌بینی به صورت جداگانه در نظر گرفته نمی‌شود. میزان قدر مطلق همبستگی پیرسون متغیرهای پیش‌بین نهایی نسبت به متغیر هدف تا آستانه ۰٫۵۳ بوده است(جدول۱). مجموعه داده نهایی در ابعاد ۵۰۰،۰۰۰*۳۴ جهت مراحل بعدی کار مهیا شد.

[1]Exploratory Data Analysis

۳-۲ پیکربندی مدل شواهد چارچوب DeepSkill

نظر به مطالعات گسترده در ادبیات مسئله و عملکرد مناسب شبکه‌های عصبی عمیق در مسائل مشابه، این پژوهش فرضیه استفاده از یادگیری عمیق به منظور پیکربندی مدل شواهد DeepSkill را در دستور کار قرار داد. بدین منظور از کتابخانه TensorFlow نسخه ۲,۴ و سیستمی با پردازنده Intel-i7-4710-HQ، پردازنده گرافیکی Nvidia GTX960 و حافظه اصلی ۱۶ گیگابایتی استفاده شد. پس از بررسی و تحلیل پیکربندی‌های مختلف از شبکه‌ای با ۵ لایه با تابع فعال‌سازی ReLU و در لایه خروجی با تابع فعالساز linear (با توجه به نیاز تخمین رگرسیونی) استفاده شد. همچنین از MSE[۱] برای تابع هزینه، و adam به عنوان الگوریتم بهینه‌ساز استفاده شد.

خروجی این شبکه یک مقدار بین صفر تا یک است(متغیر هدف "رده‌بندی تیمی" در بازه ۰ تا ۱ نرمال‌سازی شده بوده است). در این آزمایش تعریف میزان مهارت هر تیم به این صورت تعبیر خواهد شد: "احتمال قرارگیری تیم n در مقام اول" ؛ لذا هر اندازه متغیر رده‌بندی به صورت دقیق‌تری اندازه‌گیری شود، به معنای اندازه‌گیری دقیق‌تر میزان مهارت خواهد بود. نظر به پیوسته بودن متغیر هدف، خروجی به دست آمده بر اساس سه معیار MAE[۲]، MSE و همبستگی پیرسون(r)[۳]مورد ارزیابی قرار گرفت. روند یادگیری DeepSkill در تخمین مهارت تیم‌ها با دو معیار Loss(MSE) و MAE در شکل ۲ نشان داده شده است. همانطور که مشاهده می‌شود با افزایش تعداد دوره‌ها(Epoch) میزان loss برای آموزش و صحت‌سنجی به صورت نمایی در حال کاهش است که نشان دهنده بهره‌بری بالای مدل طراحی شده و افزایش سطح دقت این مدل با افزایش Epochها است. با مقایسه انجام شده بین مقادیر تخمین زده شده‌ی مهارتی و مقادیر واقعی، به ترتیب برای سه معیار MSE، MAE و r برابر با ۰,۰۲ ، ۰,۱ و ۰,۹۳ بوده است. MAE عبارت است از مقدار میانگین انحراف نتایج تخمین زده شده از مقادیر واقعی متغیر هدف. مقدار ۰,۱ برای این معیار خطای بسیار کم خروجی DeepSkill را نشان می‌دهد. معیار r بیانگر میزان همبستگی پیرسون متغیر رتبه‌بندی پیشبینی شده و متغیر رتبه‌بندی واقعی است که هرچقدر این مقدار بالاتر (این مقدار می‌تواند بین صفر تا یک متغیر باشد)باشد نشان دهنده تطابق بهتر مقادیر تخمین زده شده با مقادیر واقعی بوده است. مقدار این معیار برای نتایج خروجی DeepSkill برابر ۰,۹۳ بوده است که خطای بسیار کم توزیع خروجی شبکه استفاده شده را نشان می‌دهد.



شکل ۲ روند کاهشی Loss و MAE با افزایش تعداد تکرارهای آموزش شبکه

[۱] Mean Square Error
[۲] Mean Absolute Error
[۳] Pearson Correlation

۴-نتیجه‌گیری

تامین رضایت بازیکنان یکی از ارکان اساسی موفقیت بازی‌های رایانه‌ای در بازار است. در این پژوهش تلاش شد با ارائه چارچوبی با نام DeepSkill روشی برای سنجش مناسب سطوح مهارتی تیم‌های شرکت کننده در بازی که از یکی از چالش‌های اصلی در تامین رضایت بازیکنان در بازی‌های برخط است پوشش داده شود. با استفاده از این چارچوب و به منظور قرار دادن بازیکنان در کانال جریان، می‌توان تیم‌هایی با سطوح مهارتی متناسبی را در مقابل هم قرار داد. سه سوال اصلی در ابتدای این پژوهش مد نظر قرار گرفت. در پاسخ به سوال اول این پژوهش با بهره‌گیری از ادبیات مسئله مهارت بازیکنان در بازی‌های بچنا را در چهار مقوله اصلی دسته‌بندی نمود که با توجه به دقت‌های به دست آمده در بخش ارزیابی، در نظر گرفته شدن تمامی این ابعاد در سنجش مهارت در بازی‌های بچنا نسبت به رویکردهایی چون Trueskill می‌تواند تاثیر بسیار مثبتی داشته باشد. در پاسخ به سوال دوم، DeepSkill نیازمند تعریف هیچگونه الگوی خبرگی(مبنای اندازه‌گیری مهارت) طراحی شده توسط عامل انسانی نبوده و این چارچوب می‌تواند در تمامی بازی‌هایی با ماهیت غیر قابل پیش‌بینی جهت سنجش مهارت تیم‌ها استفاده شود. و در نهایت در پاسخ به سوال سوم، DeepSkill نشان داد با وجود نرمال‌سازی متغیرهای عملکردی بازیکنان بر اساس زمان حضور آن‌ها در بازی(مستقل‌سازی متغیرهای عملکردی از زمان)، تخمین مهارت تیمی با دقت مناسبی به دست آمده است.

مراجع

[1] S. Reis, L. P. Reis, and N. Lau, "Player engagement enhancement with video games," in *New Knowledge in Information Systems and Technologies*, Á. Rocha, H. Adeli, L. P. Reis, and S. Costanzo, Eds. Cham: Springer International Publishing, 2019, pp. 263–272.

[2] J. Schell, *The Art of Game Design: A Book of Lenses*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2008.

[3] M. Csikszentmihalyi, *Beyond Boredom and Anxiety*, ser. The Jossey-Bass behavioral science series. Jossey-Bass Publishers, 1975. [Online]. Available: https://-books.google.com/books?id=afdGAAAAMAAJ

[4] ——, *Finding flow: The psychology of engagement with everyday life.* Basic Books, 1997.

[5] G. D. Ellis, J. E. Voelkl, and C. Morris, "Measurement and analysis issues with explanation of variance in daily experience using the flow model," *Journal of leisure research*, vol. 26, no. 4, pp. 337–356, 1994.

[6] M. Qian and K. R. Clark, "EnglishGame-based learning and 21st century skills: A review of recent research," *EnglishComputers in Human Behavior*, vol. 63, pp. 50–58, 2016. [Online]. Available: www.scopus.com

[7] D. W. Shaffer, R. Halverson, K. R. Squire, and J. P. Gee, "Video games and the future of learning. wcer working paper no. 2005-4." *Wisconsin Center for Education Research (NJ1)*, 2005.

[8] R. Herbrich, T. Minka, and T. Graepel, "Trueskillâ„¢: A bayesian skill rating system," in *Proceedings of the 19th International Conference on Neural Information Processing Systems*, ser. NIPS'06. Cambridge, MA, USA: MIT Press, 2006, p. ۵۶۹□□"۵۷۶.

[9] T. Minka, R. Cleven, and Y. Zaykov, "Trueskill 2: An improved bayesian skill rating system," Microsoft, Tech. Rep. MSR-TR-2018-8, March 2018. [Online]. Available: https://www.microsoft.com/en-us/research/publication/trueskill-2-improved-bayesian-skill-rating-system/

[10] S. Omidshafiei, C. Papadimitriou, and G. Piliouras, "Î±-rank: Multi-agent evaluation by evolutionl," *Scientific Reports*, vol. 9, no. 9937, pp. 337–356, 2019.

[11] R. Bartle, *Designing Virtual Worlds*. New Riders Games, 2003.

[12] W. Min, M. Frankosky, B. Mott, J. Rowe, A. Smith, E. Wiebe, K. Boyer, and J. Lester, "Deepstealth: Game-based learning stealth assessment with deep neural networks," *IEEE Transactions on Learning Technologies*, vol. 13, no. 2, pp. 312–325, 2020, cited By 4. [Online]. Available: https://www.scopus.com/inward/-record.uri?eid=2-s2.0-85067620624&doi=10.1109%2fTLT.2019.2922356&partnerID=40&md5=5080e89755d184b1d407b656ebb23aa4

[13] R. Levy, "Dynamic bayesian network modeling of game-based diagnostic assessments," *Multivariate Behavioral Research*, vol. 54, no. 6, pp. 771–794, 2019, cited By 3. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063880588&doi=10.1080%2f00273171.2019.1590794&partnerID=40&md5=d451bcc7837755087041c9bbfd745f15

[14] R. PelÃ¡nek, "Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques," *User Modeling and User-Adapted Interaction*, vol. 27, no. 3-5, pp. 313–350, 2017, cited By 34. [Online]. Available:

https://www.scopus.com/inward/record.uri?eid=2-s2.0-85019189406&doi=10.1007%2fs11257-017-9193-2&partnerID=40&md5=a15c9607b132d517353c2301dc5d6246

[15] K. Stoeffler, Y. Rosen, M. Bolsinova, and A. A. von Davier, "Gamified performance assessment of collaborative problem solving skills," *Computers in Human Behavior*, vol. 104, p. 106036, 2020. [Online]. Available: http://-www.sciencedirect.com/science/article/pii/S0747563219302146

[16] V. J. Shute, "Stealth assessment in computer-based games to support learning," *Computer games and instruction*, vol. 55, no. 2, pp. 503–524, 2011.

[17] V. J. Shute, L. Wang, S. Greiff, W. Zhao, and G. Moore, "Measuring problem solving skills via stealth assessment in an engaging video game," *Computers in Human Behavior*, vol. 63, pp. 106–117, 2016.

[18] V. J. Shute and B. Emihovich, "Assessing problem-solving skills in game-based immersive environments," *Second Handbook of Information Technology in Primary and Secondary Education*, pp. 1–14, 2018.

[19] M. Binkley, O. Erstad, J. Herman, S. Raizen, M. Ripley, M. Miller-Ricci, and M. Rumble, *Defining Twenty-First Century Skills*, 10 2011, pp. 17–66.

[20] M. McCreery, P. Schrader, and S. Krach, "Navigating massively multiplayer online games: Evaluating 21st century skills for learning within virtual environments," *Journal of Educational Computing Research*, vol. 44, pp. 473 – 493, 01 2011.

[21] S. Webb, "Avatar culture: Narrative, power and identity in virtual world environments," *Information*, vol. 4, 01 2001.