

بناام خدا



# محاسبات جمعی: حل مساله با استفاده از جمع سپاری

**Crowd Computing: Problem Solving Using Crowd-Sourcing**

تهیه شده توسط جعفر محمدی

بهمن ۱۳۸۹

## چکیده

در این مستند به بررسی پیشنهاد شبکه‌های جمعی به عنوان ابزاری برای توسعه سیستم‌های هوشمند مصنوعی خواهیم پرداخت. شبکه‌های جمعی در سالیان اخیر به عنوان بستر مناسبی برای فعالیت‌های جمعی افراد در راستای برآورده ساختن اهداف مختلف مطرح شده‌اند. واژه‌های «جمع‌سپاری» و «هوش جمعی» نیز که به این فعالیت‌های گروهی دلالت دارند، جزو واژه‌های محبوب، به ویژه در دنیای اینترنت شده‌اند. فراگیر شدن این مفاهیم و نتایج شگرفی که در بسیاری از موارد بوجود آورده‌اند، این باور را در بسیاری از کارشناسان بوجود آورده است که تجارت و در حالت کلی‌تر فعالیت‌های آینده از جنس «جمعی» خواهد بود.

این رویکرد جدید تا کنون نقش فعالی در زمینه‌های تجارت و مدیریت باز کرده است، ولی کاربردهای آن در حوزه هوش مصنوعی بسیار اندک بوده است. مهم‌ترین دلیل آن را می‌توان پیچیدگی بیشتر این کاربردها دانست. ارائه چارچوب مناسبی برای بکارگیری این شبکه‌ها در راستای فعالیت‌های هوشمند می‌تواند کمک شایانی به این موضوع بکند.

در بسیاری از کاربردهای هوش مصنوعی، وظیفه مورد نظر برای هوشمندسازی می‌تواند به تعداد زیادی زیر وظیفه شکسته شود. این زیر وظیفه‌ها می‌توانند توسط اعضای یک شبکه جمعی حل شوند. تجمیع پاسخ‌های ارائه شده توسط اعضای می‌تواند منجر به ارائه پاسخ کلی وظیفه شود. در کاربردهای هوش مصنوعی این ریز وظایف فراتر از نظردهی یا جمع‌آوری گزارش و غیره بوده و می‌بایست پاسخی درست برای یک مساله واقعی (ولی با اندازه کوچک) ارائه شود. در نتیجه چندین مساله جدید در شبکه‌های جمعی پدیدار خواهند شد. مسائلی نظیر اینکه چه مسائلی می‌توانند با این رویکرد حل شوند؟ چه اعضای از شبکه برای پاسخ دادن به کدام ریز وظایف مناسب‌ترند؟ چقدر می‌توان به یک پاسخ ارائه شده توسط یک عضو اعتماد نمود؟ هر ریز وظیفه باید چگونه و توسط چند نفر حل شوند تا اطمینان حاصل شود که وظیفه نهایی دارای حداقل سطح معینی از اعتبار باشد؟ نحوه مدیریت و سازماندهی ریز وظایفی که مستقل از هم نبوده و دارای وابستگی می‌باشند، باید به چه صورت باشد؟ و ...

در فصل اول این مستند به بررسی «چرایی» روی آوردن به این رهیافت می‌پردازیم و بیان می‌کنیم که چرا این رهیافت می‌تواند در حل مسائلی موفق باشد که تا کنون نه انسان‌ها و نه ماشین‌ها قادر به حل آنها نبوده‌اند. در فصل دوم به بیان چارچوب پیشنهادی جهت اجرای سناریوی مطرح شده فوق پرداخته و مهم‌ترین مشکلات موجود در این چارچوب را شرح می‌دهیم. در فصل‌های آتی نیز به بررسی هر کدام از این مشکلات و نحوه فائق آمدن بر آنها خواهیم پرداخت.

**کلمات کلیدی:** شبکه جمعی، جمع‌سپاری، مساله ذهن و بدن، فعالیت غیرالگوریتمی، خروجی اشتباه یا نویزی، بازه اطمینان.

# فهرست مطالب

|    |       |                           |
|----|-------|---------------------------|
| ۲  | ۱     | مقدمه                     |
| ۲  | ۱.۱   | محاسبه‌گرایی و مشکلات آن  |
| ۸  | ۱.۲   | دیدگاه‌های فلسفی معاصر    |
| ۹  | ۱.۳   | مدل‌های جدید تجاری        |
| ۱۱ | ۱.۴   | الگوهای محاسباتی          |
| ۱۲ | ۱.۵   | جمع‌بندی و نتیجه‌گیری     |
| ۱۳ | ۲     | حل مساله با شبکه‌های جمعی |
| ۱۹ | ۳     | مرور ادبیات موضوع         |
| ۱۹ | ۳.۱   | نحوه استفاده              |
| ۲۰ | ۳.۲   | توزیع و تخصیص مسائل       |
| ۲۰ | ۱.۳.۲ | استفاده مستقیم            |
| ۲۱ | ۲.۳.۲ | مدل ثابت و هزینه‌بندی     |
| ۲۳ | ۳.۳.۲ | یادگیری فعال              |
| ۲۵ | ۳.۳   | اعتبارسنجی                |
| ۲۵ | ۱.۳.۳ | پیش‌پردازش                |
| ۲۶ | ۲.۳.۳ | تکنیک‌های اعتبارسنجی      |
| ۲۷ | ۳.۳.۳ | نحوه بکارگیری             |
| ۳۱ | ۳.۴   | سایر مباحث مرتبط          |
| ۳۱ | ۱.۳.۴ | تصحیح بایاس کاربران       |

# فصل ۱

## مقدمه

در این پیشنهاد پژوهشی می‌خواهیم به چرایی و چگونگی ایجاد سیستم‌های هوشمند از طریق شبکه‌های جمعی<sup>۱</sup> بپردازیم. در این فصل به عنوان مقدمه این پژوهش می‌خواهیم به بررسی چرایی این موضوع بپردازیم، اینکه «چرا رهیافت شبکه‌های جمعی می‌تواند رهیافت موثری برای تولید سیستم‌های هوشمند باشد»؟

برای پاسخ‌گویی به این سؤال از چهار دید مختلف به واکاوی آن می‌پردازیم. ابتدا به محدودیت‌های محاسبه‌گرایی<sup>۲</sup> به عنوان رویکرد اصلی سیستم‌های هوشمند فعلی پرداخته و ناتوانی‌های ذاتی این رویکرد را در مباحث مربوط به هوش مصنوعی بیان می‌کنیم. در این قسمت همچنین مساله آگاهی<sup>۳</sup> را نیز به عنوان یکی از مسائل مطرح روز بررسی خواهیم کرد. سپس دیدگاه‌های اتصال‌گرایی<sup>۴</sup>، نواخته‌گرایی<sup>۵</sup> و اطلاعات‌گرایی را به عنوان دیدگاه‌های مطرح در تبیین مسائل مغز و ذهن معرفی کرده و ارتباط رویکرد پیشنهادی را با این دیدگاه‌ها مورد بررسی قرار می‌دهیم. پس از آن رویکرد جمع‌سپاری<sup>۶</sup> را، به عنوان یکی از مدل‌های موثر در آینده تولید و تجارت، معرفی کرده و تطابق رویکرد پیشنهادی را با این مدل مورد بررسی قرار داده و بیان می‌کنیم که این رویکرد چگونه می‌تواند در فضای فعالیت‌های آینده منشاء پیدایش نسل جدیدی از سیستم‌های هوشمند باشد. و در انتها نیز به بررسی الگوهای<sup>۷</sup> محاسبات پرداخته و بیان می‌کنیم که روند محاسبات به سمت محاسبات جمعی انسانی می‌باشد.

### ۱.۱ محاسبه‌گرایی و مشکلات آن

در این بخش ابتدا محاسبه‌گرایی را به عنوان رویکرد اصلی سیستم‌های هوشمند فعلی معرفی کرده، سپس به بیان محدودیت‌های آن می‌پردازیم.

قسمت اول این محدودیت‌ها، ذاتی بوده و قابل برطرف شدن نیستند. قضایای ناتمامیت گودل<sup>۸</sup> اولین آشکارساز این محدودیت‌ها در حوزه علم حساب بودند. مساله توقف تورینگ<sup>۹</sup> و حالت تعمیم یافته آن (قضیه رایس<sup>۱۰</sup>) با الهام از این قضایا، ناتمامیت‌های حوزه علوم کامپیوتر را به خوبی نشان دادند. چیتین<sup>۱۱</sup> نیز با ارائه ثابت‌های امگا مرزهای این ناتمامیت را بسیار گسترده‌تر نمود.

در کنار این محدودیت‌های ذاتی، مشکلات بغرنج دیگری نیز وجود دارند که حل‌ناپذیری آن‌ها تا کنون به صورت علی‌الاصول ثابت نشده است، ولی تلاش‌های انجام گرفته برای آن‌ها نیز چندان امیدبخش نبوده‌اند و انتظار هم نمی‌رود که در آینده‌ای نزدیک بتوان راه‌حل‌های موثری برای آن‌ها ارائه نمود. مساله آگاهی

<sup>۱</sup>Collective Networks

<sup>۲</sup>Computationalism

<sup>۳</sup>Consciousness

<sup>۴</sup>Connectionism

<sup>۵</sup>Emergentism

<sup>۶</sup>Crowdsourcing

<sup>۷</sup>Paradigm

<sup>۸</sup>Gödel's incompleteness theorems

<sup>۹</sup>Turing's Halting Problem

<sup>۱۰</sup>Rice's Theorem

<sup>۱۱</sup>Chaitin

و مساله پایگاه دانش حس عام<sup>۱۲</sup> نمونه‌ای از این مسائل هستند که در این بخش به بررسی آن‌ها می‌پردازیم. در انتهای این بخش نیز به دسته‌بندی مسائلی می‌پردازیم که یا رویکردهای محاسبه‌گرایی به طور ذاتی قادر به حل آن‌ها نیستند، یا جزو مسائل اساسی حل نشده فعلی می‌باشند.

**محاسبه‌گرایی** بیان می‌کند که مغز محاسبات پیچیده‌ای را بر روی نمادهای مختلفی انجام داده و از این طریق فرآیندهای رفتاری و حالات ذهنی را بوجود می‌آورد. به بیان ساده‌تر این دیدگاه بیان می‌دارد که مغز انسان اساساً یک کامپیوتر است. این دیدگاه اساس ماشین‌های فعلی و بخش اعظمی از هوش مصنوعی را شامل می‌شود. مناقشه برانگیزترین بخش این تئوری نیز مساله آگاهی است، اینکه بیان می‌شود که محاسبات می‌تواند منجر به تولید آگاهی شود.

سیستم شناختی انسان‌ها شاید شبیه یک ماشین محاسباتی باشد ولی قطعاً دو تفاوت اساسی با یک ماشین مصنوع دارد، یکی اینکه تغییرات آن در طی زمان با روش‌های غیر قابل محاسبه‌ای صورت پذیرفته است، و دیگری اینکه می‌تواند اشتباه کند [۳۹].

ایرادات زیادی به محاسبه‌گرایی وارد است (ولی هیچ کدام از آن‌ها قطعی نیستند). به عنوان مثال این ایرادات در چند دسته مشکلات ریاضی‌وار، مشکلات بازنمایی دانش، قصدمندی، آگاهی و شناخت جاسازی شده یا برخاسته<sup>۱۳</sup> در مقاله [۳۹] به خوبی تشریح شده‌اند. به عنوان نمونه‌های دیگری از این مقالات می‌توان به [۲۴] [۲۵] [۲۶] [۲۷] [۲۸] اشاره نمود. در اینجا برخی از مشکلات مرتبط‌تر را با توضیحاتی که برای مبحث جاری مفید هستند، ارائه می‌دهیم.

**قضایای ناتمامیت گودل** در سال ۱۹۳۱ توسط کورت گودل ارائه شد. هدف اصلی این قضایا نشان دادن عدم امکان یافتن اصول موضوعه‌ای سازگار برای کل ریاضیات (مشهور به برنامه هیلبرت<sup>۱۴</sup>) بود، ولی حاصل آنها نتایج دیگری را به همراه داشت. قضایای ناتمامیت گودل بیان می‌کنند که عبارات درستی در حوزه اعداد طبیعی وجود دارند که درستی آنها قابل اثبات نیست. یا به عبارتی هویت‌هایی در ریاضیات وجود دارند که اگرچه از اصول نظریه اعداد طبیعی تبعیت می‌کنند، اما رفتاری متفاوت از این اعداد دارند. مراحل اثبات این قضایا<sup>۱۵</sup> عبارتند از [۷۲]:

۱. ساخت یک فرمول حسابی  $G$ ، معادل حکم « $G$ »: فرمول  $G$  قابل اثبات نیست».
۲. بیان اینکه  $G$  قابل اثبات است اگر و فقط اگر  $\neg G$  قابل اثبات باشد. و این یعنی ناسازگار بودن منطقی حساب یا غیر قابل اثبات بودن (تصمیم‌ناپذیر بودن)  $G$ .
۳. نشان دادن درست بودن مفهومی  $G$ ، علیرغم غیر قابل اثبات بودن آن با استنتاج صوری از اصول حساب.
۴. ناکامل بودن حساب بدلیل هم درست و هم تصمیم‌ناپذیر بودن  $G$ .
۵. اثبات اینکه «اگر حساب سازگار باشد، ناکامل است» و در نهایت نشان دادن اینکه «سازگاری حساب قابل اثبات نیست».

در اینجا، از ارائه اثبات این مراحل صرف‌نظر می‌کنیم (برای این منظور می‌توانید به [۷۲] مراجعه نمایید). برای بدست آوردن شهودی از نحوه اثبات این قضایا، در بخش بعد، اثبات مساله توقف تورینگ را ارائه می‌کنیم که از قضایای ناتمامیت گودل الهام گرفته و اثبات آن نیز از لحاظ مفهومی بسیار نزدیک به آنهاست.

این مراحل نشان دهنده این واقعیت هستند که همه حقایق حساب را نمی‌توان از اصول متعارف آن استنتاج کرد. حساب بصورت ذاتی ناکامل است. گودل پا را از این نیز فراتر گذاشته و نشان داد که حساب بطور ذاتی ناکامل است و حتی اگر اصولی نیز به گونه‌ای به اصول موضوعه آن اضافه شوند که فرمول درست  $G$  بتواند از مجموعه اصول جدید استنتاج شود، باز هم می‌توان فرمول درست ولی تصمیم‌ناپذیر دیگری را برای نشان دادن ناکامل بودن آن ارائه داد.

اهمیت ارائه این قضایا در این است که حقایق را بیان می‌کنند که با دیدگاه‌های قبلی موجود در مورد ریاضیات تناقض دارند. مهم‌ترین وجه آن نیز محدودیت بنیادینی<sup>۱۶</sup> است که به توان روش اصولی وارد می‌آید. در همین راستا مسائلی را می‌توان مثال زد که از مجموعه اصول حساب قابل استنتاج و اثبات نیستند و مثال نقضی نیز برای آنها ارائه نشده است. به عنوان نمونه‌هایی از این مسائل می‌توان به حدس گولدباخ<sup>۱۷</sup> (که بیان می‌کند هر عدد زوج مجموع دو عدد اول است) و فرضیه ریمن (که بیان می‌کند که بخش حقیقی غیر صفر خروجی تابع  $\zeta$  ریمن به ازای هر ورودی برابر ۰.۵ است) اشاره نمود.

مهم‌ترین ایرادی که با دیدی عملکردگرایانه به این قضایا وارد می‌شود این است که عبارتی که گودل اثبات‌های خود را مبتنی بر آن انجام داده است، عبارتی پیچیده و بی‌کاربرد است، در نتیجه نمی‌توان اعتبار چندانی برای این قضایا قایل شد. این اعتراض به دو دلیل وارد نیست. یکی اینکه در سال‌های اخیر عبارات

<sup>۱۲</sup>Common Sense Knowledge Base Problem

<sup>۱۳</sup>embodied or embedded cognition

<sup>۱۴</sup>Hilbert's program

<sup>۱۵</sup> این مراحل از نسخه قابل فهم‌تری از این قضایا - که توسط آلفرد تارسکی ارائه شده‌اند - اقتباس شده‌اند.

<sup>۱۶</sup> این محدودیت‌ها به این دلیل بنیادین هستند، چون با اضافه نمودن اصول دیگری نیز به اصول موضوعه، مشکل مطرح شده کماکان وجود خواهد داشت.

<sup>۱۷</sup>Goldbach's conjecture

ساده و کاربردی‌ای نیز ارائه شده‌اند که می‌توانند جایگزین عبارت مورد استفاده توسط گودل شوند [۱۵]. دلیل دیگر نیز اینکه این قضایا الهام‌دهنده کارهای بسیار اساسی بعدی از جمله مساله توقف تورینگ، قضیه رایس، ثابت‌های امگا و ... بوده‌اند.

ایراد دیگری که توسط مک‌کارتی به قضیه گودل وارد شده است [۱۶] این است که گودل در برهانش تئوری مجموعه‌های مناسبی را بکار نمی‌گیرد. مثلاً اگر وی از تئوری مجموعه‌های Zermelo-Fraenkel [۱۳] استفاده می‌کرد، به راحتی قادر بود قضیه گودل را اثبات کند. ایرادی که به این ایراد وارد است این است که تئوری مجموعه‌های ZF طوری طراحی شده است که دارای اصول موضوعه‌ای باشد که از همان ابتدا مانع بیان گزاره‌هایی نظیر گزاره بکار رفته در برهان گودل شود. به نظر می‌رسد که این کار به نحوی پاک کردن صورت مساله باشد. در ضمن این راه حل قابل توسعه به زمینه‌های دیگری که از تئوری مجموعه‌ها استفاده نمی‌کنند (نظیر مساله توقف، قضیه رایس و ثوابت امگا) نیست.

نکته مهمی که در مورد قضیه گودل باید ذکر شود این است که از قضیه گودل این نتیجه بآس‌آور بر نمی‌آید که ذهن انسان محدودیت ذاتی دارد، بلکه نتیجه این است که ذهن بشر بصورت کامل قابل فرموله کردن نیست [۱۷].

**مساله توقف تورینگ** توسط آلن تورینگ در سال ۱۹۳۶ ارائه شد. همانگونه که دیدیم، عبارتی که گودل در اثبات خود بکار می‌برد، واقعیت خاصی را در مورد خود آن عبارت مورد سوال قرار می‌دهد. تورینگ با الهام از این واقعیت، مساله توقف را پایه‌ریزی کرد<sup>۱۸</sup>. وی ثابت کرد که هیچ ماشین تورینگی نمی‌تواند در مورد توقف یا عدم توقف ماشین تورینگ دیگری اظهار نظر کند [۱۱] (مساله توقف یک دستگاه تورینگ تصمیم‌ناپذیر است). این نتیجه بعدها توسط Rice به قضیه‌ای با همین نام تعمیم یافت که بیان می‌کند که «هر خاصیت غیربدیهی در باره زبانی که توسط یک ماشین تورینگ پذیرفته می‌شود، تصمیم‌ناپذیر است» [۶]. یا به عبارتی ساده‌تر «هر سوال غیر بدیهی در باره رفتار یا خروجی یک ماشین تورینگ تصمیم‌ناپذیر است».

در ادامه این بخش، نسخه قابل فهم‌تری از اثبات مساله توقف را بصورت خلاصه ارائه می‌دهیم. اثبات مساله توقف تورینگ برای این منظور از برهان خلف استفاده می‌شود:

الف) فرض می‌کنیم مساله توقف‌پذیری یک ماشین تورینگ تصمیم‌پذیر باشد و ماشین تورینگ  $h$  وجود داشته باشد که بتواند در مورد توقف‌پذیری ماشین تورینگ  $p$  با ورودی  $i$  بصورت زیر اظهار نظر کند:

- $h(p,i)=true$  یعنی اینکه  $p$  با ورودی  $i$  متوقف می‌شود.
- $h(p,i)=false$  یعنی اینکه  $p$  با ورودی  $i$  متوقف نمی‌شود.

ب) تابع  $t(i)$  را بصورت زیر تعریف می‌کنیم:

```
Program t(input i)
  if h(i,i)=0
    stop;
  else
    loop forever;
```

پ) می‌خواهیم بررسی کنیم که ماشین تورینگ  $t$  با ورودی  $t$  متوقف می‌شود یا خیر؟

- اگر  $t(t)$  متوقف شود، آنگاه طبق الف)  $h(t,t) = true$  خواهد شد و طبق ب) می‌بینیم که  $t$  هرگز متوقف نخواهد شد.
- اگر  $t(t)$  متوقف نشود، آنگاه طبق الف)  $h(t,t) = false$  خواهد شد و طبق ب) ماشین تورینگ  $t$  باید متوقف شود.

در هر دو حالت فوق به تناقض می‌رسیم، پس طبق برهان خلف، فرض در نظر گرفته شده الف) نادرست بوده و مساله تصمیم‌پذیری یک ماشین تورینگ تصمیم‌ناپذیر است.

دلیل اصلی اهمیت ماشین تورینگ در تز Church-Turing نهفته است. این تز بیان می‌کند که یک تابع قابل محاسبه است اگر و فقط اگر بوسیله ماشین تورینگ قابل محاسبه باشد [۱۹] [۲۰]. یا هر عمل قابل محاسبه‌ای (الگوریتم‌واری) توسط یک ماشین تورینگ قابل انجام است. با توجه به اینکه کامپیوترهای

<sup>۱۸</sup> هم قضایای گودل و هم مساله توقف تورینگ، هر دو، دارای منشا الهامی واحدی هستند: ایده Diagonal Slash که کانتور برای اثبات نا شمارا بودن مجموعه اعداد حقیقی بکار برد [۱۵].  
<sup>۱۹</sup> ماشین تورینگ  $i$  یعنی آمین ماشین تورینگ در مجموعه شمارای همه ماشین‌های تورینگ (برای اثبات شمارا بودن مجموعه همه ماشین‌های تورینگ مثلاً به [۱۵] و [۴۰] مراجعه نمایند).

امروزی فقط قادر به انجام اعمال الگوریتمی هستند، پس از این تز نتیجه می‌شود که قدرت محاسباتی یک ابر کامپیوتر امروزی با قدرت ماشین تورینگ برابر است [۱۵]. این تز بدون اثبات دارای پذیرشی تقریباً جهانی است. با این حال اخیراً تلاش‌هایی نیز برای اثبات آن انجام گرفته است که به عنوان نمونه می‌توان به [۲۱]، [۲۲] و [۲۳] اشاره نمود.

از این قضایا نتیجه می‌شود که اظهار نظر غیر بدیهی در مورد رفتار یا عملکرد یک الگوریتم (ماشین تورینگ) یک فعالیت غیرالگوریتمی است. در نتیجه از یک کامپیوتر که ساختار آن بر این اساس پی‌ریزی شده است که بتواند یک فعالیت الگوریتمی را انجام دهد، نمی‌توان انتظار انجام بسیاری از فعالیت‌های لازم برای هوشمندی را داشت. این نتیجه محدودیت بسیار شدیدی را از کامپیوترهای امروزی نشان می‌دهد.

مک‌کارتی در [۱۶] معتقد است که هوش مصنوعی و علوم کامپیوتر دارای محدودیت ذاتی نیستند، چون منطق قابلیت نمایش همه آن مواردی را دارد که به عنوان محدودیت برای هوش مصنوعی مطرح هستند. وی معتقد است که در حال حاضر ضعف از الگوریتم‌ها نیست، از داده‌ها و دانش مورد نیاز است که به خوبی توسط انسان‌ها تامین نمی‌شود. ایرادی که به این ادعا وارد است این است که مک‌کارتی بخش درک مساله را که باید آن هم توسط کامپیوتر انجام شود، به انسان واگذار کرده است که چندان منطقی به نظر نمی‌رسد. در این مورد در بخش‌های بعد بیشتر صحبت خواهیم کرد.

شواهد دیگری نیز برای نشان دادن غیرالگوریتمی بودن فعالیت‌های انسانی وجود دارد. البته بدلیل اینکه اثبات محکمی برای این شواهد نمی‌تواند ارائه شود، در اینجا خیلی به آنها نمی‌پردازیم. تفکر، غیرالگوریتمی است؟ یکی از این شواهد نشان می‌دهد که تفکر عملی غیر الگوریتمی است. در تفکرات انسان‌ها کلمات جایی ندارند و آنها فقط برای بیان اتفاق‌هایی که رخ می‌دهند، از کلمات استفاده می‌کنند. یک انسان هیچ وقت نمی‌تواند تمام آنچه که در ذهن‌اش اتفاق می‌افتد را بیان کند. این شهود و توجه به این نکته که بیان الگوریتمی با کلمات اتفاق می‌افتد، می‌تواند نشان دهنده این امر باشد که تفکر انسان غیر الگوریتمی است.

**ثابت‌های امگا** مبحث دیگری هستند که با بهره‌گیری از مساله توقف در این حوزه مطرح شده‌اند. قبل از شروع این مبحث، به یادآوری تعریف محاسبه‌ناپذیری اعداد حقیقی می‌پردازیم. تعریف محاسبه‌ناپذیری اعداد بخش اعشاری اعداد حقیقی‌ای، نظیر  $\pi$  یا  $\sqrt{2}$  می‌تواند توسط یک الگوریتم تولید شود. به این اعداد محاسبه‌پذیر<sup>۲۰</sup> گفته می‌شود. در مقابل این دسته، اعداد محاسبه‌ناپذیر هستند که رقم‌های آنها توسط روش‌های الگوریتمی قابل محاسبه نیستند. به عنوان نمونه‌ای از این دسته، می‌توان به عدد دودویی‌ای اشاره نمود که صفر یا یک بودن  $\Omega$  رقمش معادل توقف یا عدم توقف  $\Omega$  امین ماشین تورینگ در مجموعه شماری همه ماشین‌های تورینگ باشد.

مجموعه تمامی ماشین‌های تورینگ شمارا هستند [۱۵]. با فرض اینکه هر ماشین تورینگ یک خروجی حقیقی تولید کند، ماشین‌های تورینگ فقط قادر به تولید تعداد شماری اعداد حقیقی، به عنوان خروجی، هستند. از آنجایی که تعداد کل اعداد حقیقی نامشمارا هستند، پس تعدادی عدد حقیقی هستند که نمی‌توانند خروجی هیچ ماشین تورینگی باشند (محاسبه‌ناپذیر هستند) اعداد محاسبه‌ناپذیر نقش مهمی را در نمایش ناتمامیت علم بازی می‌کنند. نمونه‌ای از این اعداد ثابت‌های چیتین<sup>۲۱</sup> (امگا) هستند [۹].

سر منشأ ثابت امگا نیز به مساله توقف تورینگ بر می‌گردد. ارتباط ثابت امگا و مساله توقف تورینگ احتمال توقف یک ماشین تورینگ که به تصادف از بین همه ماشین‌های تورینگ انتخاب شود، چقدر است؟ چیتین این احتمال را  $\Omega$  (امگا) نامید. در ابتدا محاسبه  $\Omega$  و حتی فرموله کردن مساله آن غیر ممکن می‌نمود، ولی این کار با یافته ارائه شده در مقاله [۱۲] مبنی بر اینکه می‌توان عملیات هر ماشین تورینگ را به یک معادله دیوفانتی ترجمه کرد، امکان‌پذیر شد. در این مقاله همچنین نشان داده شد که توقف هر ماشین تورینگ معادل داشتن جواب طبیعی معادله دیوفانتی متناظر آن ماشین می‌باشد (بدیهی است که مساله داشتن یا نداشتن جواب طبیعی یک معادله دیوفانتی، تصمیم‌ناپذیر می‌باشد). چیتین با استفاده از معادلات دیوفانتی توانست مساله محاسبه  $\Omega$  را فرموله کرده و سپس نشان دهد که  $\Omega$  محاسبه‌پذیر نیست. نشان داده شده است که در صورت امکان محاسبه  $\Omega$ ، می‌توان مسائل نظیر حدس گلدباخ و فرضیه ریمان را نیز حل نمود.

کار به همین جا ختم نمی‌شود. کامپیوترهای خداگونه هم احتمال توقف دارند. چیتین سوپر امگاها را نیز مطرح نمود که درجه تصادفی بودن آنها بسیار بالاتر از  $\Omega$  است. اگر یک کامپیوتر خداگونه<sup>۲۲</sup> وجود داشته باشد که قادر به حل مساله توقف بوده و بتواند  $\Omega$  را محاسبه کند، باز هم یک احتمال توقف ( $\Omega'$ ) برای این کامپیوتر وجود خواهد داشت. این روند کماکان ادامه خواهد داشت و احتمال‌های توقف برای کامپیوترهای قوی‌تر و قوی‌تر نیز مطرح خواهد بود. لازم به ذکر است که این امگاها صرفاً ثابت‌های ریاضی بی‌معنی نیستند. به عنوان مثال مفهوم سه امگای اول از رشته امگاها، بصورت زیر توسط چیتین ارائه شد:

- $\Omega$ : احتمال توقف یک ماشین تورینگ به تصادف انتخاب شده.
- $\Omega'$ : احتمال تولید تعدادی متناهی خروجی توسط یک سری محاسبات نامتناهی.
- $\Omega''$ : احتمال fail کردن یک سری محاسبات نامتناهی در تولید خروجی.

<sup>۲۰</sup> computable

<sup>۲۱</sup> Chaitin's constants

<sup>۲۲</sup> omnipotent

وجود این رشته امگاها نشان می‌دهد که هرگز نمی‌توان کامپیوتری ساخت که بتواند همه چیز را محاسبه کند.

**مساله آگاهی** یکی از مسائل باز و مطرح فعلی در دنیای هوش مصنوعی می‌باشد. برای مطرح کردن این مساله ابتدا به تعریف آگاهی در فلسفه ذهن می‌پردازیم. تعریف یکسان و واحدی برای آگاهی، که دارای پذیرش جهانی باشد، وجود ندارد. آگاهی برای یک موجود در ساده‌ترین تعریف خود، عبارت است از داشتن حالت ذهنی «حس آن موجود بودن». معتبرترین تعریف نیز برای آگاهی در مساله سخت آگاهی<sup>۲۳</sup> نهفته است که توسط چالمرز در سال ۱۹۹۵ ارائه شده است. مساله سخت آگاهی این است که چگونه می‌توان مکانیزم فیزیکی را با تجربه ذهنی هوشیاری تطبیق دهیم. چگونه فرآیندهای فیزیکی در مغز به تجارب ذهنی مرتبط می‌شوند؟ بسیاری بر این باورند که تجارب ذهنی از پایه‌های فیزیکی برمی‌خیزند، در حالی که هیچ تبیینی برای چگونگی این برخاستن وجود ندارد. [۲۹] تعریف مطرح دیگری برای آگاهی تئوری تفکر یا درک مرتبه بالاتر<sup>۲۴</sup> می‌باشد که توسط روزنتال و درتسکی ارائه شده‌اند. طبق این تئوری آگاهی عبارت است از بازنمایی حالات ذهنی یا رفتاری یک سطح در سطحی بالاتر.

توجه داشته باشیم که باید بین آگاهی و همبسته‌های عصبی آگاهی<sup>۲۵</sup> تفاوت قائل شویم. به عنوان مثال آگاهی علی‌الاصول قابل اندازه‌گیری نیست ولی می‌توان ادعا نمود که همبسته‌های عصبی آن را می‌توان اندازه گرفت [۳۰].

دلیل اهمیت آگاهی در این است که دانشمندانی نظیر جان سرل یا پن‌روز معتقدند که تعریف هوشمندی بدون ذکر آگاهی ناقص است و حضور توأم تفکر (رفتار) و آگاهی است که هوشمندی را می‌سازد [۱۵]. البته فارغ از تعریف هوشمندی، مساله آگاهی مصنوعی یا بوجود آوردن عامل‌های آگاه، اخیراً توجه بسیاری از دانشمندان را به خود جلب کرده است.

مطرح‌ترین دیدگاهی که در تبیین آگاهی و برخاستن آن از پایه‌های فیزیکی وجود دارد از دیدگاه‌های اتصال‌گرایی و نوحاسته‌گرایی بهره می‌برد. در این دیدگاه‌ها ادعا می‌شود که آگاهی ویژگی هیچ کدام از نرون‌های مغز به تنهایی نیست، بلکه یک ویژگی کل‌نگرانه است که از ارتباط و همکاری بین نرون‌ها نشأت می‌گیرد، در حالی که سیستم به حد معینی از پیچیدگی برسد. با بهره‌گیری از این دیدگاه مقاله [۲۴] به بررسی امکان آگاهی مصنوعی پرداخته است. در این مقاله حدی از پیچیدگی که در آن ارتباط بین نرون‌ها منجر به بوجود آمدن آگاهی می‌شود، معادل پیچیدگی (تعداد و میزان ارتباطات نرون‌ها) مغز انسان در نظر گرفته شده است. و بیان شده است که پیاده‌سازی این تعداد نرون و ارتباطات بین آن‌ها در صورتی محقق خواهد شد که حافظه کامپیوترها به سطح معینی برسد. پیش‌بینی این مقاله طبق قوانینی نظیر قانون مور این است که در حدود بیست سال آینده به این سطح از حافظه در کامپیوترها خواهیم رسید و متعاقباً نیز بوجود آورد آگاهی مصنوعی امکان‌پذیر خواهد شد. پیش‌بینی‌های مشابهی توسط برخی دیگر از نویسندگان نیز ارائه شده است. از جمله [۳۵] [۳۶] و [۳۷].

نماینده مناسبی برای دیدگاه فیزیک‌گرایی در این حوزه نیز دانیل دنت می‌باش. دنت معتقد است که ساخت یک مغز مصنوعی اساساً به همان اندازه امکان‌پذیر است که ساخت یک قلب مصنوعی، فقط این کار بسیار بسیار سخت‌تر است. از طرفی وی آگاهی را محصول مستقیم مغز می‌داند و معتقد است که اگر یک کپی سلول به سلول از یک نفر ساخته شود، احمقانه! است که فکر کنیم که آن کپی نتواند همانند اصل خود آگاهی داشته باشد! [۳۲]

با این وجود بسیاری دیگر از دانشمندان نیز هنوز یا هیچ تبیین قابل قبولی را نیافته‌اند یا معتقدند که یافتن چنین تبیینی با ابزارهای امروزی امکان‌پذیر نیست.

البته تلاش‌هایی نیز برای شبیه‌سازی آگاهی در سطح سیستم‌های کامپیوتری صورت گرفته است که تلاش می‌کنند سیستم‌هایی را بسازند که آگاهی نداشته باشد، ولی قادر به شبیه‌سازی آن باشد. پن‌روز چنین تلاش‌های را از پایه نادرست می‌پندارد و معتقد است که نمی‌توان موجودی ساخت که خودآگاه نباشد ولی از دید یک کاربر بیرونی معادل یک موجود خودآگاه عمل (فکر) کند. چون خودآگاهی ضرورتاً خود را در برخی از اعمال (یا تفکرات) آشکار می‌سازد.

**مساله پایگاه دانش حس عام** بیان می‌دارد که چگونه می‌توان تمامی حقایقی که یک آدم می‌داند را ذخیره کرده و به آن دسترسی پیدا کرد [۴۳]؟ این مساله در حوزه هوش مصنوعی توسط ماروین مینسکی در سال ۱۹۹۲ مطرح شد. مینسکی در [۸] به بررسی وضعیت هوش مصنوعی می‌پردازد. دلایل پیشرفت‌های کند در زمینه‌های هوش مصنوعی یکی از دلایلی که وی برای پیشرفت‌های کند در زمینه‌هایی نظیر درک زبان طبیعی بیان می‌کند این است که کامپیوترها به معانی کلمات و اشیا، به آن صورتی که انسان‌ها می‌فهمند، دسترسی ندارند. وی برای تشریح این واقعیت واژه «ریسمان» را مثال می‌زند. یک انسان به آسانی می‌فهمد که می‌توان با ریسمان چیزی را «کشید»، ولی نمی‌شود چیزی را «هل داد». با یک ریسمان می‌شود چیزی را «بسته‌بندی کرد» ولی نمی‌توان آن را «خورد» و ... یک نوجوان می‌تواند در چند دقیقه صدها کاربرد یا عدم کاربرد ریسمان را برای شما بازگو کند ولی یک کامپیوتر قادر به انجام حتی بخشی از این کار نیست. حال کافی است به جای ریسمان صدها هزار کلمه و مفهوم دیگری را در نظر بگیرید که می‌شناسید. اگر کامپیوتری بخواهد زبان طبیعی را درک کند باید چنین درکی از تمامی این کلمات و مفاهیم و ارتباطات بین آنها داشته باشد، که ندارد.

این امر برای سایر کاربردها (به جز درک زبان طبیعی) نیز صادق است. مثلاً کاربرد بینایی ماشین و مساله تشخیص شی صندلی را در نظر بگیرید. یک انسان

<sup>۲۳</sup>The hard problem of consciousness

<sup>۲۴</sup>Higher Order Thought or Higher Order Perception Theory

<sup>۲۵</sup>Neural Correlates of Consciousness



برای درک صندلی نیازی ندارد که حتما شی متعارفی را ببیند که دارای چهار پایه و یک پشتی است، وی با دیدن یک صندلی با هر تعداد پایه و یا داشتن و یا نداشتن پشتی یا ... قادر است صندلی را تشخیص دهد. وی برای این منظور از درک صندلی و کاربردهای و همچنین ارتباط آن با سایر اشیای دنیا که در ذهن وی شکل گرفته است، بهره می‌برد. اگر کامپیوتر نیز بخواهد که بینایی داشته باشد باید چنین درکی از تمامی اشیا و ارتباط بین آنها داشته باشد، که ندارد.

مینسکی از این مشکل تحت عنوان «نبود پایگاه دانش حس عام» نام می‌برد. این مشکل دارای وجوه مختلف زیر است [۴۱] [۴۲]:

۱. ساخت چنین پایگاه دانشی از لحاظ حجم کاری مورد نیاز عملا غیر ممکن است.
۲. با فرض انجام‌پذیر بودن این کار، پیدا کردن روشی برای نمایش<sup>۲۶</sup> مناسب این دانش بسیار مشکل است.
۳. بروز رسانی تغییرات در دانش و دنبال کردن تمامی حقایقی که بوسیله تغییرات تأثیر می‌پذیرند، مشکل است.
۴. طراحی و توسعه سیستمی که بتواند از این پایگاه دانش استفاده کند، اگر ناممکن نباشد، حداقل بسیار سخت است.

مینسکی همچنین بیان می‌کند که برنامه‌های کامپیوتری وجود دارند که بتوانند شطرنج بازی کنند، یا یک تصویر متنی را به متن تبدیل کنند، ولی هیچ برنامه شطرنجی وجود ندارد که بتواند کار تبدیل متن را انجام دهد، یا اینکه هیچ برنامه تبدیل متنی نمی‌تواند شطرنج بازی کند. وی مهم‌ترین دلیل این امر را عدم استفاده برنامه‌های مختلف از یک پایگاه دانش حس عام می‌داند.

واقعیتی که وجود دارد این است که مساله حس عام هنوز هم به عنوان یک مشکل اساسی در راه پیشرفت هوش مصنوعی وجود دارد و تلاش‌ها در این زمینه چندان موفق نبوده است.

**مسائل ناتمام علوم کامپیوتر.** در این بخش به این موضوع پرداختیم که هوش مصنوعی دارای محدودیت‌های بنیادین و ذاتی است. این محدودیت‌ها فقط دارای جنبه‌های تئوری نیستند و کاربردهای عملی زیادی از آنها متأثر هستند. در اینجا می‌خواهیم مهم‌ترین مشکلات موجود در این زمینه را دسته‌بندی نموده و به ارائه نمونه‌های عملی از مسائل موجود در این دسته بپردازیم.

این دسته‌بندی شامل دو دسته کلی می‌باشد. دسته اول مسائلی که علی‌الاصول توسط تکنولوژی فعلی قابل حل نیستند، و دسته دوم مسائلی که ثابت نمی‌شود که غیر قابل حل‌اند، ولی آفاق روشنی برای حل آنها وجود ندارد. در دسته‌بندی زیر منظور از یک عمل سطح بالاتر در مورد یک «چیز» مواردی نظیر درک، تحلیل، اظهار نظر یا قضاوت در مورد آن «چیز»، مد نظر می‌باشد. مسائلی که علی‌الاصول توسط هوش مصنوعی قابل حل نیستند عبارتند از:

۱. **مسائلی که فرآیند انجام آنها غیر الگوریتمی است.** برخی از این مسائل غیر الگوریتمی به راحتی توسط انسان‌ها قابل حل هستند و برخی دیگر برای انسان‌ها نیز ناشناخته‌اند. مسائلی نظیر مساله به ظاهر ساده غیربازگشتی تصمیم‌گیری در مورد قابل پوشش بودن یا نبودن یک سطح مسطح اقلیدسی با استفاده از کاشی‌های پوشاننده<sup>۲۷</sup> [۱۵]، تبدیل بدون خطای گفتار یا تصاویر متنی به متن، تگ کردن و حاشیه‌نویسی تصاویر و سایر مستندات چندرسانه‌ای، تحلیل عواطف موجود در یک متن، تصویر یا صدا، مشکل شرایط آب‌وهوایی کره زمین و ... نمونه مسائلی از این دسته می‌باشند.
۲. **فعالیت‌هایی که نیاز به یک عمل سطح بالاتر در مورد رفتار یا عملکرد خود آن ماشین دارند.** به عنوان درک گیر افتادن در حلقه بی‌نهایت پس از مدت زمان قابل قبولی و توقف فعالیت جاری، نمونه‌ای از کارهایی است که انسان به راحتی می‌تواند انجام دهد، ولی یک ماشین هرگز قادر به انجام آن نیست، زیرا نمی‌تواند بصورت غیر استنباطی<sup>۲۸</sup> به یک حالت سطح پائین تر توجه کند<sup>۲۹</sup>.
۳. **فعالیت‌هایی که نیاز به یک عمل سطح بالاتر در مورد رفتار یا عملکرد ماشین‌های دیگر دارند.** نمونه خیلی واضحی از این دسته مسائل، قضاوت در مورد مناسب بودن یا نبودن یک الگوریتم کامپیوتری برای حل یک مساله مشخص می‌باشد. تست بتای نرم‌افزارها هم نمونه خوب دیگری می‌باشد. هیچ کامپیوتری نمی‌تواند تست‌های غیر بدیهی را در مورد یک نرم‌افزار بتا انجام دهد.
۴. **فعالیت‌هایی که نیاز به یک عمل سطح بالاتر در مورد رفتار یا عملکرد انسان‌ها دارند.** ارزیابی رفتار انسان‌ها، قضاوت‌های حقوقی، ارزیابی اثرات هنری و ... نمونه‌ای از این دسته مسائل می‌باشند.

<sup>۲۶</sup> representation

<sup>۲۷</sup> nonrecursive tiling

<sup>۲۸</sup> noninferential

<sup>۲۹</sup> روزنتال نشان می‌دهد که توجه استنباطی به یک حالت سطح پائین تر مستلزم دوری بودن است.

## ۵. فعالیت‌هایی که نیاز به یک عمل سطح بالاتر در مورد مفاهیم و حقایق کلی موجود در جهان دارند...

مسائلی که ثابت نمی‌شود که غیر قابل حل هستند ولی افق روشنی برای حل آنها در آینده نزدیک وجود ندارد عبارتند از:

۱. **مسائلی که برای انجام آنها نیاز به پایگاه دانش حس عام می‌باشد.** مسائلی نظیر تشخیص اشیا و فهم زبان طبیعی نمونه‌هایی از این دسته می‌باشند...

۲. **فعالیت‌هایی که نیاز به درک محیط یا مساله دارند.** در کامپیوترهای امروزی، یک عامل بیرونی وجود دارد که اطلاعات مورد نیاز یک مساله را پردازش کرده و آنها را به فرمی در می‌آورد که توسط کامپیوتر قابل دریافت و پردازش باشند (یک مدل انتزاعی و در بسیاری از موارد ناکامل از مساله را برای کامپیوتر فراهم می‌کند که با درک خود مساله توسط کامپیوتر تفاوتی بنیادین دارد). سپس کامپیوتر اقدام به حل مساله با توجه به آن ورودی‌ها می‌کند. البته در این مرحله نیز کامپیوترها محاسباتی را که به آنها واگذار می‌شود انجام می‌دهند، بدون اینکه درکی از آنها داشته باشند (در واقع کامپیوترها مقداری از درک برنامه‌نویس‌هایشان را بکار می‌گیرند. برنامه‌نویس مساله را درک کرده و فرموله می‌کند و برای حل آن الگوریتم ساده یا حتی مبتنی بر یادگیری ماشین ارائه می‌دهد. منتها به دلیل سرعت، دقت و حافظه کامپیوترها، اجرای الگوریتم را به آنها واگذار می‌کند [۱۷]). کامپیوتری وجود ندارد که بتواند مساله‌ای را درک کرده و نسبت به حل آن اقدام نماید. قدرت الگوریتمی هر کامپیوتر وابسته به قدرت تحلیل و طراحی برنامه نویسی دارد که برنامه وی را می‌نویسد.

## ۱.۲ دیدگاه‌های فلسفی معاصر

در اینجا به بررسی چند دیدگاه معاصر می‌پردازیم که در تبیین مسائل ذهن و بدن موفق‌تر از سایرین عمل کرده‌اند. دلیل ذکر این دیدگاه‌ها در این است که هر کدام به قسمی تأیید کننده رویکرد ارائه شده می‌باشند. این دیدگاه‌ها عبارتند از اتصال‌گرایی، نوحاسته‌گرایی و اطلاعات‌گرایی. در ادامه به بررسی این دیدگاه‌های می‌پردازیم.

**دیدگاه‌های اتصال‌گرایی و نوحاسته‌گرایی.** بدلیل نزدیکی دو دیدگاه فوق از جنبه‌ای که می‌خواهیم به آن بپردازیم، این دو دیدگاه را با هم بررسی می‌کنیم. دیدگاه اتصال‌گرایی دیدگاهی در زمینه نحوه تولید حالات ذهنی می‌باشد که تلاش می‌کند تا پدیده‌های رفتاری یا ذهنی را به عنوان فرآیندهایی ناشی از عملکرد شبکه بهم پیوسته‌ای از بخش‌های ساده مدل کند. مهم‌ترین مدل ارائه شده در این زمینه شبکه‌های عصبی (طبیعی و مصنوعی) می‌باشند. اتصال‌گرایان معتقدند که ارتباطات موجود مابین تعداد زیادی از نرون‌های فیزیکی باعث بوجود آمدن فرآیندهای ذهنی (غیر فیزیکی) می‌شوند. دیدگاه اتصال‌گرایی به مادی‌گرایان کمک می‌کند تا بتوانند چگونگی وجود فرآیندهای ذهنی را توجیه کنند.

روش‌های هوش گروهی محاسباتی<sup>۳۰</sup> نظیر کلونی مورچه‌ها<sup>۳۱</sup> مبحث دیگری هستند که نزدیکی معنایی بالایی به دیدگاه اتصال‌گرایی دارند. به عنوان مثال مقاله [۷] عملکرد مغز را با استفاد از مفهوم کلونی مورچه‌ها توضیح می‌دهد. کلونی مورچه‌ها یک مبحث داغ در سالیان اخیر بوده است. روح کلی این مبحث و سایر الگوریتم‌ها و ایده‌های الهام گرفته شده از آن این است که از همکاری مورچه‌ها برای حل مسائل پیچیده بهینه‌سازی استفاده شود. هر مورچه به تنهایی دارای هوش و توانمندی‌های محدودی می‌باشد ولی جمعی از مورچه‌ها در یک کلونی دارای رفتار جمعی بسیار پیچیده و هوشمندی می‌باشند. بدیهی است که هیچ کدام از مورچه‌ها به تنهایی قادر نخواهند بود کل پروژه (مثلاً فضای مسیریابی) را تصور کرده و عملکرد مورچه‌ها و نتیجه نهایی را دریابند یا به تنهایی به آن دست یابند.

از طرفی در دیدگاه نوحاسته‌گرایی، ویژگی‌های برخاسته تعریف می‌شوند که طبق تعریف نمی‌توانند توسط اجزای یک سیستم تعیین شوند بلکه کل سیستم است که آنها را می‌سازد. مثلاً ویژگی کوتاهترین مسیر در کلونی مورچه‌ها. در این حالت حتی تعریف مساله کوتاهترین مسیر و چگونگی روبرو شدن برای حل آن توسط اجزای سیستم قابل درک نیست [۳۳]. نوحاسته‌گرایان ذهن را چیزی مجزا از بدن نمی‌دانند، بلکه آن را برخاسته از بدن می‌دانند. نوحاسته‌گرایی قوی ادعا می‌کند که آگاهی نیز می‌تواند از بدن برخیزد.

ارتباط این دو دیدگاه با رویکرد پیشنهادی در این است که اولاً هوش گروهی محاسباتی می‌تواند از دنیای کلونی مورچه‌ها و دسته‌های پرندگان مهاجر به شبکه‌های جمعی آدم‌ها گسترش داده شود. و دوم اینکه اگر ویژگی‌های ذهنی و حتی آگاهی بتواند از نرون‌های فیزیکی برخیزد، پس می‌توان انتظار داشت که در یک شبکه جمعی مفاهیم سطح بالاتری از ارتباط و همکاری بین انسان‌های آگاه برخیزد!

<sup>۳۰</sup> computational swarm intelligence

<sup>۳۱</sup> ant colony

**دیدگاه اطلاعات‌گرایی** در این دیدگاه اطلاعات نقش اساسی در تبیین عملکرد سیستم‌ها دارد. سیستم‌ها در این دیدگاه بر اساس اینکه عملکردشان بر اساس انتقال اطلاعات در بین اجزایشان باشد یا اینکه فقط بر اساس نیروهای کنش و واکنش بین آنها، به دسته‌های گوناگونی تقسیم می‌شوند. از طرفی در سیستم‌هایی که عملکردشان بر اساس تبادل اطلاعات بین اجزایشان صورت می‌گیرد بر اساس اینکه اطلاعات آنها اصیل باشد یا بازتابی از اطلاعات اصیل، باز به دو زیر دسته تقسیم می‌شوند. سیستم‌های مصنوع (یا بخش‌های یک سیستم مصنوع) دارای اطلاعات اصیل نیستند و بازتاب اطلاعاتی هستند که ما در آنها گذاشته‌ایم (با تعیین شرایط اولیه یا تنظیم نحوه اندازه‌گیری‌ها) [۳۰]. از طرفی ادعا می‌شود که فقط یک سیستم با تعامل اطلاعات اصیل است که می‌تواند آگاهی داشته باشد. این کار با استفاده از مفهوم قصدمندی<sup>۳۳</sup> انجام می‌پذیرد. از طرفی آگاهی و قصدمندی دو مفهوم کاملاً در هم تنیده هستند و از طرفی دیگر فقط موجودات و سیستم‌های بیولوژیکی فی‌نفسه قصدمند هستند (قصدمندی مصنوعات، آئینه‌ای از قصدمندی سازنده آنها در راستای هدف مشخصی بوده و مانند قصدمندی موجودات بیولوژیکی اصیل نیست) [۳۰].

ملیح از همین استدلال بهره می‌گیرد و ادعا می‌کند که اگر تمامی نوروهای مغز یک انسان با چیپ‌های سیلیکونی هم‌ارز آنها عوض شوند، موجود بوجود آمده دارای آگاهی نخواهد بود، چون تعامل بین نوروها بر پایه تبادل اطلاعات اصیل بوده و teleonomical می‌باشد، در حالی که تعامل بین چیپ‌های سیلیکونی teleological بوده و نمی‌تواند بصورت طبیعی وجود داشته باشد، مگر با استفاده از شانس ذاتی [۳۰]!

دیدگاه اطلاعات‌گرایی بر تفاوت بین سیستم‌های شامل موجودات بیولوژیکی و مصنوعی تاکید کرده و بسیاری از ویژگی سیستم‌های هوشمند را فقط برای این سیستم‌ها قائل است. بر اساس این دیدگاه یک سیستم شبکه جمعی دارای توانمندی‌های ذاتی بسیار بیشتری از سیستم‌های هوشمند صرفاً ماشینی می‌باشد.

## ۱.۳ مدل‌های جدید تجاری

در این بخش به بررسی شبکه‌های جمعی<sup>۳۴</sup> می‌پردازیم که به عنوان ابزار قدرتمندی برای انجام فعالیت‌های جدید مطرح شده‌اند. برخی از متخصصین دنیای اقتصاد و مدیریت، معتقدند که انجام فعالیت‌ها بر اساس شبکه‌های جمعی، آینده دنیای کسب و کار را دچار تحول می‌کنند [۴۴][۴۵].

امروزه، زندگی ما در شبکه‌های جمعی و اجتماعی مختلف سپری می‌شود. هر فرد عضو یک یا چند شبکه می‌باشد. بسیاری از این شبکه‌ها نداشت دیجیتالی داشته و به خوبی قابل مطالعه هستند. در این مطالعات می‌توان تاثیر فرد و شبکه را متقابلاً بر روی هم بررسی نمود.

هر فردی در هر شبکه، دارای قدرت اطلاعاتی مشخصی می‌باشد. قدرت اطلاعاتی شبکه نیز در این است که در آن، حتی اگر عضوی قادر باشد فعالیت هر کدام از اعضای دیگر آن شبکه را انجام دهد، قادر به انجام برآیند فعالیت همه اعضای آن شبکه نخواهد بود. یعنی دانش قابل حصول از اعضای یک شبکه به صورت شگرفی بیشتر از دانش تکین اعضای آن شبکه می‌باشد.

در ادامه ابتدا مفهوم جمع‌سپاری را تعریف نموده و کاربردهای فعلی آن را بر خواهیم شمرد. سپس ارتباط جمع‌سپاری و شبکه‌های جمعی را با موضوع پیشنهادی را بیان خواهیم نمود.

جمع‌سپاری<sup>۳۴</sup> ترکیبی معنایی از دو عبارت خرد جمعی<sup>۳۵</sup> و برون‌سپاری<sup>۳۶</sup> می‌باشد. این اصطلاح نخستین بار در سال ۲۰۰۶ توسط روزنامه‌نگاری به نام Jeff Howe در مقاله‌ای در نشریه Wired معرفی شد [۸]. وی در سال ۲۰۰۸ نیز کتابی را با همین عنوان منتشر نمود [۲].

طبق تعریف ویکی‌پدیا<sup>۳۷</sup> جمع‌سپاری عبارت است از عمل برون‌سپاری وظایف بوسیله یک کارفرما به گروه بزرگی از مردم یا یک انجمن (یک جمع) از طریق یک فراخوان. طبق توصیف ویکی‌پدیا، جمع‌سپاری یک مدل توزیع‌شده حل مساله و تولید می‌باشد. ویکی‌پدیا تفاوت بین جمع‌سپاری و برون‌سپاری را در این می‌داند که در جمع‌سپاری بر خلاف برون‌سپاری، مساله بین گروهی از کاربران ناشناس پخش می‌شود نه یک فرد مشخص شناخته شده. تفاوت بین جمع‌سپاری و متن‌باز هم در این است که در فعالیت‌های متن‌باز فعالیت‌های داوطلبانه‌ای هستند که توسط یک فرد از همان جامعه آغاز می‌شوند و سایر اعضا نیز به شکل داوطلبانه به تکمیل آن می‌پردازند، ولی در فعالیت‌های جمع‌سپاری یک مشتری است که اقدام به شروع یک فعالیت می‌کند و گروه‌هایی از اعضای شبکه بصورت موازی به حل آن فعالیت اقدام می‌کنند. تفاوت اساسی دیگری که بین آن دو وجود دارد، انگیزه‌های شرکت کنندگان در آن فعالیت‌ها می‌باشد. در حالت کلی هر فعالیت متن‌باز می‌تواند به عنوان یک فعالیت جمع‌سپاری تلقی شود، ولی عکس این قضیه صادق نیست.

<sup>۳۳</sup>Intentionality

<sup>۳۳</sup> شبکه‌های جمعی (collective networks) شبکه‌هایی هستند که افراد موجود در آنها [بر خلاف شبکه‌های اجتماعی] دارای ارتباطات مستقیم بسیار اندکی هستند.

<sup>۳۴</sup>Crowdsourcing

<sup>۳۵</sup>The wisdom of crowd

<sup>۳۶</sup>Outsourcing

<sup>۳۷</sup><http://en.wikipedia.org/wiki/Crowdsourcing>

سرعت رشد جمع‌سپاری بعد از معرفی ایده اولیه آن بسیار سریع بوده و شبکه‌های بسیار زیادی با هدف جمع‌سپاری بوجود آمده‌اند. مهم‌ترین کاربردهای این شبکه‌ها را می‌توان بصورت زیر دسته‌بندی نمود<sup>۳۸</sup>:

**خلق و ایجاد.** در این کاربرد از اعضای شبکه خواسته می‌شود که محتوا، محصول یا مفهومی را خلق کنند. از مهم‌ترین نمونه‌های این دسته می‌توان به نرم‌افزارهای متن‌باز، ویکی‌پدیا، شبکه‌هایی نظیر **Cambrian House** برای طراحی محصولات نرم‌افزاری تجاری یا **CrowdSpirit** برای طراحی و اجرای محصولات الکترونیکی ارزان قیمت تجاری اشاره نمود.

**نیروهای آماده بکار.** در این کاربرد نیز هدف خلق و ایجاد است، با این تفاوت که هدف‌های معین و از قبل مشخص شده‌ای وجود ندارد، بلکه اعضای شبکه همانند نیروهای آماده بکاری هستند که منتظرند تا درخواستی به سیستم وارد شود تا بنا به آن درخواست بلافاصله خلق و ایجاد خود را شروع کنند. به عنوان نمونه‌ای از این دسته می‌توان به شبکه **Rent A Coder** اشاره نمود که دارای تعداد بسیار زیادی طراح و توسعه دهنده نرم‌افزاری به عنوان عضو است که از آنها برای طراحی و توسعه پروژه‌های نرم‌افزاری جذب شده بهره می‌گیرد.

**تحقیق و توسعه.** شبکه‌های این کاربرد اقدام به جذب متخصصین و کارشناسان به عنوان عضو می‌کنند. سپس از صنایع و شرکت‌های بزرگ، درخواست می‌کنند تا نیازهایشان را در جهت رفع مشکلات تحقیق و توسعه خود یا ایجاد نوآوری و خلاقیت در فرآیندهای این شبکه اعلام کنند. با اعلام نیازها به شبکه، این نیازها به اعضا اعلام می‌شود. در صورتی که راهکاری از طریق یک یا تیمی از اعضا برای این پروژه‌ها پیشنهاد شود، این راهکار بصورت یک پیشنهاد به شرکت متقاضی ارائه می‌شود. در اصل این شبکه‌ها شبیه بخش تحقیق و توسعه برای تمامی شرکت‌هایی هستند که از خدمات آنها استفاده می‌کنند. با این تفاوت که دارای تعداد بسیار زیادی متخصص بوده و سربار مالی برای آن شرکت ندارند. قدرت برخی از این شبکه‌ها به حدی است که به عنوان مرجعی جهت حل مشکلات تکنولوژیک شناخته می‌شوند و شرکت‌های Fortune Top 500 از مشتریان پر و پا قرص آنها می‌باشند. نمونه قوی‌ای از این شبکه‌ها، شبکه **InnoCentive** باییش از دویست هزار عضو می‌باشد که ۶۱ درصد آنها دارای مدارج پیشرفته علمی یا صنعتی هستند. مشتریان این شبکه نیز شرکت‌هایی هستند که با چالش‌های جدی روبرو هستند و توسط تیم تحقیق و توسعه شرکت خود قادر به حل این چالش‌ها نبوده یا می‌خواهند با هزینه و زمان پایین‌تری به حل مشکل‌شان دست یابند. سازمان‌های **The Economist**، **Nature**، **NASA**، **SAP**، بنیاد راکفلر و ... از مشتریان این شبکه می‌باشند.

**ایده‌پردازی و ایده پروری.** در این کاربرد اعضای شبکه به ایده‌پردازی پرداخته و سعی می‌کنند ایده‌هایی را شناسایی کنند که امکان تجاری‌سازی آنها وجود داشته باشد. جذاب‌ترین ایده‌ها پس از رای‌گیری از اعضا مشخص شده و جهت بهره‌برداری‌های مختلف مورد استفاده قرار می‌گیرند.

**سرمایه‌گذاری جمعی** در شبکه‌های سرمایه‌گذاری جمعی، اعضا به منظور پیدا کردن راه‌هایی جهت انجام سرمایه‌گذاری بر روی دارایی‌هایشان گرد هم می‌آیند. هر فردی در این شبکه می‌تواند ایده‌ای را برای سرمایه‌گذاری مطرح نماید. در صورت مقبولیت ایده بین سایر اعضا، آنها سرمایه‌هایشان را برای عملی کردن پیشنهاد ارائه شده در اختیار این عضو قرار می‌دهند. معمولاً در این شبکه‌ها روال‌های مشخصی برای نظارت‌ها، تعهدها و گارانتی‌ها وجود دارد. وجود جمع زیادی از افراد خبره در امور مالی و سرمایه‌گذاری و نظارت و نظرات سایر اعضا در این شبکه‌ها، شکست سرمایه‌گذاری را به حد بسیار پایینی می‌رساند.

**پیش‌بینی.** در این کاربرد شبکه‌های جمع‌سپاری سعی می‌کنند تا از دانش اعضای شبکه برای پیش‌بینی میزان موفقیت یک محصول، میزان محبوبیت یک طراحی یا مفهوم، پیامدهای یک اتفاق و ... استفاده کنند. این شبکه‌ها را به نحوی می‌توان وسیله‌ای برای سنجش نظرات کاربران یک محصول، سنجش افکار عمومی، جمع‌آوری افکار افراد خبره و ... دانست. نمونه‌هایی از این شبکه‌ها **Threadless** (برای پیش‌بینی اینکه کدام تی‌شرت‌ها در بازار خوب فروش می‌کنند) و **Dell Idea Storm** (برای پیش‌بینی اینکه چه ویژگی‌هایی در محصولات آتی Dell وجود داشته باشند، این شرکت در بازار موفق‌تر خواهد بود) می‌باشند.

**سازمان‌دهی.** در این شبکه‌ها، جمع‌سپاری در راستای سازمان‌دهی اطلاعات خاصی مورد استفاده قرار می‌گیرد. نمونه‌هایی از این شبکه‌ها عبارتند از: **dig** (لینک‌های وب مختلفی به این شبکه ارسال می‌شود و کاربران در امتیاز دهی به آنها شرکت می‌کنند تا لینک‌هایی که ارزش چک کردن دارند، برای سایرین مشخص نمایند).

**هوش جمعی** در شبکه‌های هوش جمعی<sup>۳۹</sup> وظایف و پروژه‌هایی با هدف انجام توسط اعضای شبکه پذیرش می‌شوند. این پروژه‌ها یا وظایف به تعداد بسیار زیادی ریز وظیفه شکسته می‌شوند، بطوری که هر ریز وظیفه توسط تعدادی از افراد شبکه قابل انجام باشد. در نهایت جواب‌های این ریز وظایف با هم ادغام شده و جواب وظیفه یا پروژه نهایی از آن ساخته می‌شود. معمولاً در این شبکه‌ها وظیفه شکست وظیفه، اعتبار سنجی پاسخ‌های ارائه شده توسط کاربران و جمع آنها را صاحبان وظایف بر عهده دارند. افراد شبکه نیز به ازای انجام ریز وظایف مالی را دریافت می‌کنند. شبکه **Mechanical Turk** آمازون معروف‌ترین شبکه از این دسته می‌باشد. به عنوان نمونه‌های دیگری از این دسته می‌توان به این شبکه‌ها اشاره نمود: **Distributed Proofreaders** (برای تصحیح متون ماشینی دیجیتالی شده)، **Google Image Labeler** (یک بازی هدفدار ساده برای لیبیل‌گذاری تصاویر جهت بازیابی محتوایی آنها)، **Mob4Hire** (برای تست برنامه‌های کاربردی تحت موبایل) و **NASA Clickworkers** (برای پردازش انسانی تصاویر مریخ).

<sup>۳۸</sup> برای اطلاعات بیشتر در این زمینه و مشاهده نمونه‌های متنوعی از هر دسته به گزارش تکنیکی [۷۳] مراجعه نمایید.

<sup>۳۹</sup> Collective Intelligence

**جمع‌سپاری و استخراج دانش.** دانش بسیار زیادی به صورت پیدا و پنهان در شبکه‌ها وجود دارد. این دانش غالباً «پراکنده» بوده یا «پنهان» می‌باشد. «پراکنده‌گی» از این لحاظ که هیچ‌کدام از افراد آن شبکه به همه آن دانش دسترسی ندارند و هر عضوی از شبکه ممکن است به بخشی از آن دسترسی داشته باشد. «پنهان» نیز از این لحاظ که ممکن بخشی از دانش که هر فرد می‌تواند به آن دسترسی داشته باشد، بصورت شسته و رفته در اختیار وی نباشد و لازم باشد که وی مقداری انرژی برای دستیابی به آن خرج کند. در بسیاری از موارد افراد علاقه‌ای به صرف انرژی برای پیدا کردن دانش پنهان اطراف‌شان ندارند، چون به این مساله واقف هستند که این دانش، فقط جزء کوچکی از کل دانش مربوطه بوده و به تنهایی برای‌شان فایده‌ای به همراه نخواهد داشت. اگر ناظر سطح بالاتری به وجود این دانش پراکنده و پنهان واقف باشد، می‌تواند برای دستیابی و بهره‌برداری از کل آن با استفاده از یک شبکه جمعی یا اجتماعی اقدام کند. مثال‌های عملی‌ای وجود دارند که نشان داده‌اند که نه تنها دستیابی به دانشی از این جنس امکان‌پذیر است، بلکه سرعت دستیابی به آن نیز با اتخاذ استراتژی مناسب، بسیار سریع می‌باشد.

بارزترین مثال موجود در این بخش مسابقه بال‌های قرمز دارپا<sup>۴۰</sup> می‌باشد. مسابقه بال‌های قرمز، نشان دهنده قدرت اطلاعاتی شبکه‌های جمعی دارپا به مناسبت چهلیمین سالگرد تولد اینترنت، مسابقه‌ای را تحت عنوان «چالش شبکه‌ای دارپا»، در اواخر سال ۲۰۰۹ به اجرا گذاشت<sup>۴۱</sup>. هدف از برگزاری این مسابقه نشان دادن نقش اینترنت و شبکه‌های اجتماعی در ارتباطات زماندار، ساخت تیم‌های مجازی با اعضای پراکنده و بسیج فوری برای حل مسایل با حوزه گسترده و بحرانی از لحاظ زمانی بود. این مسابقه همچنین نمایش مناسبی از کارایی جمع‌سپاری در شبکه‌های سنسوری بود [۲]. در این مسابقه از شرکت‌کنندگان خواسته شده بود که ده بال‌شانور هشت فوتی را که در ده نقطه ناشناس آمریکا به مدت محدود از یک زمان مشخص نمایش داده می‌شدند، را شناسایی کنند. به اولین فرد یا گروهی که موقعیت تمامی این بال‌ها را به درستی گزارش می‌داد، جایزه‌ای تعلق می‌گرفت.

مهم‌ترین معضل پیش روی شرکت‌کنندگان در این مسابقه غلبه بر تاکتیک‌های منفی رقبا بوده است. مهم‌ترین تاکتیک‌های منفی در اختیار تیم‌ها عبارت بوده است از<sup>۴۲</sup>: بر افراشتن بال‌های مشابه، ارسال گزارش‌های به ظاهر مستند ولی اشتباه برای تیم‌های رقیب توسط افراد متعدد و در نهایت از بین بردن بال‌های در دسترس جهت کاهش شانس پیدا شدن آنها توسط دیگر تیم‌ها.

حدود چهار هزار و سیصد گروه در این مسابقه شرکت کردند که در نهایت تیمی از دانشگاه MIT موفق به بردن این مسابقه شد. اگر چه تیم‌ها برای شناسایی این بال‌ها نه روز فرصت داشتند، ولی این تیم در کمتر از نه ساعت موفق به گزارش صحیح تمامی بال‌ها شد. استراتژی این تیم شکل‌دهی یک شبکه خاص منظوره (به جای استفاده از شبکه‌های عمومی نظیر Facebook، تویتر و ...) و ترغیب افراد به عضویت در این شبکه با بکار گرفتن روش‌های بازاریابی شبکه‌ای بوده است. گزارشی از استراتژی‌های غلبه بر گزارش‌های نادرست که توسط این تیم بکار گرفته شده بود، ارائه نشده است.

در این بخش مفهوم جمع‌سپاری و قدرت شبکه‌های جمعی به تفصیل مورد بررسی قرار گرفت. رویکرد پیشنهادی از شبکه‌های جمعی به عنوان بستر خود استفاده می‌کند. با توجه به توضیحات ارائه شده دستیابی به سیستم‌های هوشمند و استخراج دانش و قابلیت‌های انسانی با بکارگیری توانمندی‌های اعضای این شبکه‌ها دور از دسترس نیست.

## ۱.۴ الگوهای محاسباتی

در این بخش به بررسی الگوهای محاسباتی از ابتدا تا امروز می‌پردازیم و می‌بینیم که گرایش این الگوها به سمت محاسبات جمعی می‌باشد.

الگوهای محاسبات را می‌توان از دیدهای مختلفی دسته‌بندی کرد. یکی از این دیدها، روند تکامل آن‌ها در طی زمان می‌باشد. الگوی محاسباتی دهه ۱۹۶۰ محاسبات دسته‌ای<sup>۴۳</sup> بود که با استفاده از اتاق‌های کامپیوتر انجام می‌گرفت. در دهه ۱۹۷۰ و با ظهور mainframe و ترمینال‌ها، الگوهای محاسباتی اشتراک زمانی<sup>۴۴</sup> بوجود آمد. با ظهور کامپیوترهای شخصی در دهه ۱۹۸۰ دو الگوی فوق از بین رفته و نوبت به محاسبات رومیزی<sup>۴۵</sup> رسید. این الگو در حال حاضر، با ظهور پردازنده‌های با تعداد هسته‌های بالاتر در حال تکامل است. آخرین دسته از الگوهای محاسبات نیز محاسبات شبکه‌ای می‌باشند که در دهه ۱۹۹۰ و با ظهور شبکه‌های کامپیوتری بوجود آمدند. این دسته بیشترین تکامل و گسترش را نسبت به دسته‌های پیشین داشته‌اند و در حال حاضر نیز بیشترین تمرکز بر روی آن‌ها می‌باشد. در ادامه به بررسی تکامل الگوهای این دسته می‌پردازیم.

در این دسته، ابتدا محاسبات توزیع شده مطرح شدند که در آن‌ها یک برنامه به چند تکه شکسته شده و این تکه‌ها هم‌زمان بر روی کامپیوترهای مختلفی که از

<sup>۴۰</sup> آژانس تحقیقات پیشرفته دفاعی در ایالات متحده آمریکا

<sup>۴۱</sup> <https://networkchallenge.darpa.mil>

<sup>۴۲</sup> <http://mssv.net/2009/10/31/how-to-win-the-darpa-network-challenge/>

<sup>۴۳</sup> Batch

<sup>۴۴</sup> Time-Sharing

<sup>۴۵</sup> Desktop

طریق شبکه به هم متصل بودند، اجرا می‌شدند. سپس نوبت به محاسبات خوشه‌ای رسید که در آن‌ها کامپیوترهای با سخت‌افزار و نرم‌افزار همسان در یک زیر شبکه<sup>۴۶</sup> بصورت تنگاتنگی بهم متصل می‌شدند. یک سیستم مرکزی نیز پردازنده‌ها را بصورت موازی اجرا کرده و توزیع بار را بر روی آن‌ها طوری انجام می‌داد که تعادل بار متوازن بر روی آن‌ها برقرار باشد. نسل بعدی محاسبات توری<sup>۴۷</sup> هستند که بسیار انعطاف‌پذیرتر از محاسبات خوشه‌ای هستند. در این الگو هر کدام از کامپیوترهای مجموعه می‌تواند دارای سخت‌افزار و نرم‌افزار متفاوتی باشد و ارتباطات نیز می‌تواند بین زیر شبکه‌های مختلفی برقرار گردد (رفع محدودیت‌های سخت‌افزاری، نرم‌افزاری و مکانی). همچنین در این محاسبات گره‌های مختلف می‌توانند به مجموعه اضافه شده یا از آن حذف شوند (پویایی در مقیاس توری).

جدیدترین الگو که در حال حاضر بیشترین میزان پژوهش‌ها و سرمایه‌ها را به خود اختصاص داده است، محاسبات ابری<sup>۴۸</sup> می‌باشد. در الگوی محاسبات توری هر کدام از گره‌ها باید دارای حداقل قدرت محاسباتی مشخصی باشند تا برای کل مجموعه مفید واقع شوند، ولی در محاسبات ابری، گره‌های مختلف دارای هر قدرتی باشند سیستم مرکزی می‌تواند از توانایی‌های آن بهره بگیرد. در این محاسبات ابر به «اینترنت» اشاره داشته و نشان دهنده این است که ارتباط گره‌ها (سرویس دهندگان) و سرویس گیرندگان از طریق اینترنت (با انعطاف‌پذیری بسیار بالاتری) انجام می‌شود.

نسل نوظهور این الگوها، محاسبات جمعی<sup>۴۹</sup> می‌باشد. در این الگو بر خلاف محاسبات ابری که در آن گره‌های سرویس دهنده متعلق به گروه یا شرکت خاص هستند، منابع کامپیوتری از طریق تمامی کاربران اینترنت می‌توانند به اشتراک گذاشته شوند. هر کاربر نیز می‌تواند بابت به اشتراک‌گذاری منابع خود هزینه یا خدماتی را از سرویس گیرندگان دریافت دارد.

با بررسی این روند و پژوهش‌هایی که در حال شکل‌گیری است، می‌توان نسل آینده این الگوها را محاسبات جمعی ترکیبی (انسانی و ماشینی) دانست. در این الگوی محاسباتی، شبکه منابع مورد استفاده همانند محاسبات جمعی ماشینی می‌باشد، با این تفاوت که تعدادی گره انسانی نیز به مجموعه افزوده می‌شوند. بخشی از فعالیت‌ها، برای انجام، به ماشین‌ها سپرده شده و برخی نیز به گره‌های انسانی واگذار می‌شوند. در این الگو نیز هر کاربر بابت فعالیت‌هایی که برای مجموعه انجام می‌دهد یا منابعی که در اختیار آن قرار می‌دهد هزینه یا خدماتی را از سرویس گیرندگان دریافت می‌کند.

## ۱.۵ جمع‌بندی و نتیجه‌گیری

در این فصل به چرایی استفاده از شبکه‌های جمعی در راستای بوجود آوردن سیستم‌های هوشمند اشاره کردیم. دلیل اصلی روی آوردن به این دیدگاه این است که ماشین‌های امروزی قادر به حل مسائل بسیار ساده غیر الگوریتمی نیستند ولی انسان‌ها به راحتی می‌توانند این مسائل را حل کنند. از طرفی سرعت و دقت پردازش کامپیوترها در مقایسه با انسان‌ها بسیار بیشتر است. در نتیجه ترکیب قدرت غیر الگوریتمی مغز انسان‌ها و سرعت و دقت کامپیوترها کاملاً منطقی به نظر می‌رسد. در این فصل تبیین‌هایی از دیدگاه‌های مختلف فلسفی را نیز بیان کردیم که بر درست بودن این تصمیم دلالت داشتند. لازم به ذکر است که دلایل ذکر شده در حد تبیین می‌باشند نه اثبات، چون اصولاً ارائه اثباتی برای این منظور امکان‌پذیر نیست.

رویکردهای دیگری نیز علاوه بر این رویکرد ارائه شده‌اند، ولی هیچ‌کدام دارای پیشرفت‌های چشمگیری نبوده‌اند. به عنوان مثال پن‌روز ایده استفاده از کامپیوترهای کوانتومی را برای تطابق بیشتر با ساختار فیزیولوژی انسان مطرح می‌کند [۱۵]. ایده استفاده از کامپیوترهای کوانتومی ایرادی که به این پیشنهاد وارد است این است که کامپیوترهای کوانتومی ممکن است بتوانند برخی از مسائلی را که در کامپیوترهای معمولی در زمان نامایی قابل حل هستند را در زمان چند جمله‌ای حل کنند، ولی از لحاظ قدرت محاسباتی همانند کامپیوترهای فعلی بوده و قادر به حل مسائل غیر الگوریتمی نمی‌باشند [۱۵]. یا به عنوان مثالی دیگر و به منظور تقویت ماشین‌های تورینگ، تلاش‌هایی برای ارائه فرمولاسیونی جهت کامپیوترهای آنالوگ صورت گرفته است تا چارچوب ماشین‌های حالت انتزاعی<sup>۵۰</sup> را به مدل‌های پیوسته زمان توسعه دهد. این تلاش‌ها نیز در ابتدای مسیر خود قرار دارند [۳۸].

به عنوان آخرین بخش این فصل دورنمایی از آینده سیستم‌های هوش مصنوعی از طریق شبکه‌های جمعی را ارائه می‌دهیم. دورنمای سیستم‌های جمعی در این دورنما برنامه‌های کامپیوتری بسیار هوشمندی وجود دارند که می‌توانند فعالیت‌های بسیار زیادی را که قبلاً قادر به انجام‌شان نبودند را به خوبی انجام دهند. فعالیت‌هایی که فراتر از محدودیت‌های نسبت داده توسط قضیه رایس به آنها می‌باشد. آنها همچنین می‌توانند زیبایی را درک کنند، قضاوت کنند، اخلاقیات را درک کرده و در مورد آن اظهار نظر کنند، احساسات موجود در متون را درک کرده و بیان کنند و .... پشت پرده این برنامه‌های کامپیوتری سیستم‌های جمعی هستند که به شبکه‌های بسیار بزرگی متصلند که دارای میلیون‌ها نفر عضو هستند. هر عضو این شبکه در طی روز صدها و یا حتی چند هزار ریز وظیفه ارجاعی (الگوریتمی و غیر الگوریتمی) در زمینه‌های مختلف را به سرعت انجام داده و تحویل سیستم می‌دهد و ....

<sup>۴۶</sup> subnet

<sup>۴۷</sup> Grid

<sup>۴۸</sup> Cloud Computing

<sup>۴۹</sup> Crowd Computing

<sup>۵۰</sup> abstract state machine

## فصل ۲

# حل مساله با شبکه‌های جمعی

در فصل قبل بررسی کردیم که چرا رویکرد شبکه‌های جمعی برای حل مساله مناسب است. در این فصل به دنبال بررسی «چگونگی» این امر هستیم. چگونه می‌توان افراد و کامپیوتر را با هم مجتمع کرد به قسمی که ترکیب بوجود آمده هوشمندتر از هر دو جز تشکیل دهنده باشد؟<sup>۱</sup> چگونه می‌توان مسائلی که توسط کامپیوترها قابل حل نیستند، ولی صرفنظر از زمان لازم برای حل آنها براحتی می‌توانند توسط گروهی از انسان‌ها حل شوند، را با تکنولوژی امروزی حل نمود؟ چگونه می‌توان دانش حل مساله را که در جامعه توزیع شده است (و اخذ، انتقال و استفاده از آن در محیط جدید دارای هزینه‌های بالایی است [۱۴]) را با هزینه‌های پایینی به صورت کارا اخذ نموده و از آن بهره برد؟

همانگونه که خواهیم دید، معماری یک شبکه جمعی و اجزای مختلف آن بسیار وابسته به مساله‌ای هستند که قرار است توسط این شبکه حل شود. به همین دلیل در ابتدا سعی می‌کنیم تا یک چهارچوب کلی برای یک سیستم حل مساله با شبکه جمعی طراحی کنیم. سپس برای هر کدام از بخش‌های آن بررسی کنیم که وظیفه آن بخش چیست، برای مواجهه با مسائل مختلف، باید چه ویژگی‌هایی داشته باشد و مشکلات اساسی برای دست یافتن به وظایفش کدام است؟

برای این منظور از تحلیل سناریوی کلی مورد استفاده این شبکه‌ها بهره می‌بریم: «یک مساله در قالب این رویکرد به تعداد بسیار زیادی ریز مساله شکسته می‌شود. بخشی از ریز مسائل الگوریتمی بوده و می‌توانند توسط ماشین حل شوند. بخشی نیز غیر الگوریتمی بوده و نیاز به هوش انسانی برای حل دارند. این ریز مسائل به یک شبکه جمعی سپرده می‌شوند تا توسط اعضای آن شبکه شوند. از آنجایی که ریز مسائل این بخش غیر الگوریتمی هستند، اعتبارسنجی پاسخ‌های ارائه شده برای آنها نیز باید توسط همین شبکه انجام پذیرد. پس از اعتبارسنجی مسائل و دریافت پاسخ‌هایی با ضریب اطمینان قابل قبول، این پاسخ‌ها همراه با پاسخ‌های ارائه شده توسط ماشین برای ریز مسائل الگوریتمی، جمع شده و پاسخ نهایی مساله را بدست می‌دهند. هم‌زمان ضریب اطمینان پاسخ کلی ارائه شده نیز محاسبه شده و همراه با پاسخ نهایی ارائه می‌شود».

این سناریو، بخش‌های مختلف یک سیستم هوش جمعی را به خوبی نشان می‌دهد. در ادامه به ذکر هر کدام از این بخش‌ها پرداخته و موارد مهم مرتبط با هر کدام از آنها را ذکر می‌کنیم.

**مسائل قابل حل توسط این چارچوب.** اولین مطلبی که سناریوی فوق بیان می‌کند، این است که تنها مسائل خاصی هستند که می‌توانند توسط این رویکرد حل شوند. در این بخش به بررسی این مسائل از دو دید ماهوی و کاربردی خواهیم پرداخت. ماهوی، یعنی اینکه مسائلی که مناسب این سیستم هستند، باید چه ویژگی‌هایی داشته باشند. و کاربردی یعنی اینکه چه کاربردهای واقعی‌ای وجود دارند که دارای این ویژگی‌ها بوده، و این رویکرد قادر است پاسخی بهتر از پاسخ‌های موجود فعلی برای آنها ارائه دهد؟

یک مساله (کاربرد) مناسب برای حل توسط این رویکرد، باید دارای ویژگی‌های ساختاری زیر باشد:

- قابل شکست به تعداد زیادی ریز مساله باشد، که وابستگی کمی به هم داشته باشند (قابلیت حل ریز مسائل بصورت موازی امکان‌پذیر باشد). همچنین ریز مسائل بدیت آمده باید ایستا بوده و در طی زمان تغییر نکنند.
- بخش قابل توجهی از ریز مسائل راه حل الگوریتمی نداشته، توسط یک فرد عادی به راحتی قابل حل بوده و وابستگی به فرد پاسخ‌کننده نداشته باشند.

<sup>۱</sup>برگرفته از هدف تحقیقاتی مرکز هوش جمعی دانشگاه MIT.

همچنین یک فرد عادی تا حد قابل قبولی قادر به اعتبارسنجی پاسخ‌های ارائه شده توسط سایر افراد باشد.

- حل کل مساله توسط گروه کوچکی از افراد، دارای هزینه بالایی باشد.

- روش الگوریتم‌واری برای شکست مساله و ترکیب ریز مسائل وجود داشته باشد.

به عنوان یک مثال از جامعه هدف می‌توان به مساله بازیابی بر اساس محتوای تصویر، در سطح مفهومی اشاره نمود. در رویکرد هوش جمعی، استخراج ویژگی برای تصاویر پایگاه داده مورد نظر توسط اعضای شبکه جمعی انجام می‌پذیرد. در این کاربرد، ویژگی‌های استخراجی برای هر تصویر را می‌توان مفاهیمی در نظر گرفت که در آن تصویر وجود دارند. استخراج مفاهیم موجود در یک تصویر همانگونه که قبلاً اشاره شد، یک کار غیر الگوریتمی بوده و می‌بایست برای صدها هزار تا چند میلیون تصویر انجام پذیرد. از طرفی نیز جمع‌آوری و سازماندهی تصاویر و مفاهیم استخراجی و جستجو در این مفاهیم، کاری است که بخوبی می‌تواند از عهده یک کامپیوتر برآید. با توجه به توضیحات ارائه شده، مشاهده می‌شود که این مساله می‌تواند یک مساله هدف مناسب برای رویکرد پیشنهادی باشد.

به عنوان مثال‌های دیگری از مسائل هدف می‌توان به موارد زیر اشاره نمود: تصحیح اشتباهات و ارتقای سطح کیفی متون (از لحاظ املائی، انشایی، گرامری و افزونگی)، فیلترینگ (مثلاً اینکه آیا این محتوا برای نوجوانان مناسب است؟)، سنجش میزان مرتبط بودن نتایج برگشتی با کلید واژه‌های جستجو در خدمات جستجوگرها، تشخیص اجناس مشابه ارائه شده توسط فروشندگان مختلف در فروشگاه‌های مختلف، ساخت مجموعه داده‌های آموزشی لیبیل‌دار در کاربردهای یادگیری ماشین (مثلاً پروژه نوکیا [۴۸] برای جمع‌آوری مجموعه داده‌های مناسب، از زبان‌های مختلف، برای توسعه سرویس‌های بر پایه بازشناسی گفتار در گوشی‌های تلفن همراه)، مسائل عملی یادگیری فعال<sup>۲</sup> که فاقد لیبیل برای نمونه‌های آموزشی جدید می‌باشند، استخراج ساختارهای پروتئین از طریق بازی‌های جمعی، تحقیقات میدانی گسترده، رونویسی فایل‌های صوتی، تست بتای نرم‌افزارها، استخراج عواطف موجود در متون، تصاویر و ویدیو، کشف خطاهای رخ داده در خروجی سیستم‌های هوشمند کلاسیک نظیر برنامه‌های تبدیل تصویر به متن و ....

برای بدست آوردن شهودی از خروجی‌های قابل انتظار از یک شبکه جمعی، نگاهی به نتایج بدست آمده در [۴۶] برای کاربرد تصحیح اشتباهات و ارتقای سطح کیفی متون تهیه شده توسط نرم‌افزار Microsoft Word می‌اندازیم. نتایج آزمایش‌های انجام شده، حاکی از کاهش ۱۰ تا ۲۲ درصدی طول حجم متون در خلاصه‌سازی، تصحیح حدود ۶۷ درصدی مشکلات موجود در متن به تهایبی و حدود ۸۲ درصدی مشکلات با کمک غلطیاب خود نرم‌افزار بوده است (در این کاربرد با کمک جعبه ابزار ارائه شده توسط شبکه جمعی MTurk - که بعداً آن را معرفی می‌کنیم - استفاده از هوش جمعی به صورت یک نوار ابزار در نرم‌افزار فوق تعبیه شده و امکان تلفیق نتایج حاصل از هوش جمعی با نتایج ارائه شده توسط غلطیاب خود نرم‌افزار امکان‌پذیر شده است). البته این نتایج هزینه مالی بالایی ۱.۴۱ دلار به ازای هر پاراگراف را نیز در برداشته‌اند. در مورد هزینه زمانی نیز، با وجود اینکه زمان کلی صرف شده بسیار بالا بوده است، ولی نکات قابل تاملی در این زمینه وجود داشته‌اند. هزینه زمانی شامل دو بخش بوده است: هزینه زمانی برای جذب کاربر جهت انجام کار، و هزینه زمانی انجام خود کار. هزینه زمانی جذب، ده‌ها برابر بیشتر از هزینه زمانی انجام کار بوده است. البته اگر بپذیریم که با رشد جمع‌سپاری، اعضای شبکه‌های جمعی نیز رشد چشمگیری خواهند کرد، زمان انجام کار بسیار پایین آمده و هزینه زمانی کلی، قابل قبول خواهد بود (با در نظر گرفتن تعداد بالایی از کاربران و با احتساب اینکه پاراگراف‌های مختلف می‌توانند بصورت موازی تصحیح و خلاصه‌سازی شوند، زمان انجام کل کار برای هر متنی، در حد کمتر از ۵ دقیقه خواهد بود).

مقاله [۵۱] آزمایش‌های جالبی را به منظور مقایسه صحت فعالیت‌های انجام شده توسط افراد خبره و کاربران شبکه جمعی MTurk انجام داده است. در این مقاله برای این منظور پنج دسته فعالیت در حوزه پردازش زبان طبیعی در نظر گرفته شده است (این فعالیت‌ها عبارتند از: تحلیل عواطف موجود در متن، شباهت جملات از لحاظ معنایی، بازشناسی استلزام متنی، تشخیص رخداد‌های زمانی و تشخیص معنای کلمات مبهم). برای هر فعالیت نیز چندین مساله طراحی شده‌اند. این مسائل ابتدا برای حل به افراد خبره داده شده و سپس توسط جمعی از کاربران شبکه فوق حل شده‌اند. استراتژی‌های بکار گرفته شده برای اعتبارسنجی پاسخ‌های ارائه شده توسط کاربران و نحوه استخراج پاسخ نهایی از آن‌ها را در بخش‌های بعد مورد بررسی قرار می‌دهیم، و در اینجا فقط به ذکر نتایج بدست آمده می‌پردازیم. این نتایج بسیار قابل توجه بوده و در تمامی کاربردها به خروجی افراد خبره نزدیک شده یا از آنها پیشی گرفته‌اند. به عنوان مثال در تست تحلیل عواطف، کاربران در مورد پنج حس از شش حس (به جز حس ترس) توانستند از نظر یک فرد خبره پیشی بگیرند. در این کاربرد حداقل تعداد افراد لازم برای اینکه درستی فعالیت آنها در حل یک مساله از یک فرد خبره در حل همان مساله کمتر نباشد، بین ۲ تا ۹ نفر (برای حس‌های مختلف) متفاوت بوده است. در بررسی شباهت معنایی جملات نیز همبستگی نظرات ۱۰ نفر کاربر، تقریباً نزدیک نظر یک فرد خبره بوده است. همچنین در تمامی فعالیت‌ها، نتیجه نهایی کار با استفاده از تعداد کاربران بیشتر، بالاتر می‌رفته است (نمودار همبستگی نظرات یا درستی فعالیت بر حسب تعداد کاربران مشارکت‌کننده، صعودی بوده است، البته تا حد معینی که بستگی به مساله دارد). در این آزمایش با توجه به هزینه و زمان صرف شده، بطور متوسط ۸۴۰ ریز وظیفه در ساعت و ۱۵۰.۷ ریز وظیفه به ازای هر دلار انجام شده است.

**تولید ریز مسائل.** سناریوی ذکر شده برای یک سیستم هوش جمعی، بیان می‌کند که یک مساله برای حل توسط این رویکرد، می‌بایست به تعداد زیادی

<sup>۲</sup> Active learning



زیر مسائل شکسته شود. این زیر مسائل به طور قابل توجهی از مساله اصلی کوچکتر و ساده‌تر می‌باشند. نکته مهمی که در این زمینه وجود دارد، این است که رفتار بخش‌های مختلف سیستم خیلی وابسته به نوع مسائلی است که به آن داده می‌شوند. پس لازم است در اینجا، بررسی‌های بیشتری در مورد ریز مسائل تولیدی صورت پذیرد.

ساده‌ترین مسائل برای حل توسط سیستم‌های جمعی، مساله‌های کلاسه‌بندی طبیعی، و ساده‌ترین حالت آن یعنی مسائل دو کلاسه می‌باشند. به عنوان مثال «آیا این جمله دارای غلط املائی می‌باشد؟»، «آیا محتوای این صفحه برای کودکان مناسب است؟»، «آیا این متن حاوی حس شادی است؟» و ... جذابیت این مسائل به این دلیل است که اعتبارسنجی پاسخ‌های ارائه شده برای آنها با تکنیک رای اکثریت، به راحتی می‌تواند انجام پذیرد. در ساده‌ترین حالت کافی است یک مساله برای حل به جای یک نفر به چند نفر سپرده شده و رای مورد توافق اکثریت آنها برگشت داده شود. مسائل چند کلاسه از این لحاظ دارای پیچیدگی بیشتری و مسائل غیر کلاسه‌بندی نیز بسیار مشکل می‌باشند. زیرا اعتبارسنجی پاسخ‌های ارائه شده برای مسائل غیر کلاسه‌بندی، کار راحتی نبوده و تکنیک‌هایی نظیر رای اکثریت نیز به دلیل تعدد بسیار بالای جواب‌های ممکنه، بصورت عملی جوابگو نیستند. در این حالت یا باید راه‌حلی برای مواجهه سیستم با این گونه مسائل ارائه شود، یا مساله طوری به ریز مسائل شکسته شود، که هر کدام از ریز مسائل یک مساله کلاسه‌بندی شود. روش اول نیازمند تحقیق و پژوهش‌های جدید می‌باشد، ولی روش دوم در صورت امکان، فقط بخش شکست مساله را تحت تاثیر قرار می‌دهد.

مطلب فوق بر این واقعیت دلالت دارد که شناخت شیوه‌های شکست مساله به ریز مسائل بسیار مهم بوده و همچنین اگر تبدیل مساله به ریز مسائل، به روش‌های مختلفی امکان‌پذیر باشد، بررسی اینکه کدام روش، با توجه به سایر بخش‌های سیستم، کارتر می‌باشد، مساله بسیار مهمی است.

لازم به ذکر است که در برخی مسائل، شکست مساله فراتر از معنای متعارف آن بوده و شامل استخراج ریز مسائل می‌باشد. در این صورت می‌توان ابتدا از هوش جمعی به منظور استخراج مسائل و سپس به منظور حل آن‌ها استفاده نمود. به عنوان مثال مساله تصحیح و ارتقای سطح کیفی متون، یک مساله کلاسه‌بندی نبوده و همچنین ریز مسائل آن نیز مشخص نمی‌باشد (مشخص نیست که کدام پاراگراف‌ها یا جملات متن نیاز به تغییر دارند و باید به عنوان ریز مسائل وارد سیستم شوند). برای حل این مساله در [۴۶]، ابتدا هر کدام از پاراگراف‌ها، در یک ریز مساله به شبکه داده شده و از کاربران خواسته می‌شود تا جملاتی که نیاز به تصحیح دارند را مشخص نمایند (این مساله، یک مساله کلاسه‌بندی متعارف نمی‌باشد). در این مرحله، جملات نیازمند تصحیح مشخص شده و در مرحله بعدی، هر کدام از آن جملات، به عنوان یک ریز مساله برای تصحیح به شبکه داده می‌شود (دلیل اینکه خود پاراگراف‌ها برای تصحیح به شبکه داده نمی‌شوند، این است که اگر مسائل به اندازه کافی ریز نشوند، پاسخ‌هایی که از شبکه دریافت می‌شوند دارای کیفیت چندان بالایی نخواهند بود). تصحیح یک جمله نیز یک مساله کلاسه‌بندی نبوده و مساله سختی به شمار می‌رود. در این حالت می‌توان با تکنیک‌هایی، این مساله را به چند مساله کلاسه‌بندی تبدیل نمود. توضیحات بیشتر مربوط به این روش‌ها را در بخش‌های بعد خواهیم دید.

نحوه شکست یک مساله به ریز مسائل، وابسته به مساله بوده و عوامل بسیاری در آن دخیل هستند. در بسیاری از موارد، چگونگی شکست مساله به ریز مسائل است که تعیین می‌کند که آیا هوش جمعی برای حل آن مساله می‌تواند کارا باشد یا خیر؟ نحوه شکست مساله، جنس ریز مسائل تولیدی، موعد زمانی تحویل ریز مسائل به کاربران سیستم، پاداش در نظر گرفته شده برای حل هر کدام از ریز مسائل، مدت زمان در نظر گرفته شده تا انقضای هر ریز مساله و ... نمونه‌هایی از عواملی هستند که در میزان موفقیت‌آمیز بودن یا نبودن حل یک مساله با رویکرد شبکه‌های جمعی نقش اساسی دارند.

یکی از روش‌های مناسب برای سنجش تاثیر این عوامل، رفتارسنجی کاربران یک شبکه موجود، با آزمایش‌های کنترل شده می‌باشد. مثلاً برای سنجش تاثیر پاداش در نظر گرفته شده برای پاسخ‌دهی به مسائل، می‌توان مسائلی از کاربردهای متفاوت را، هر کدام چند بار ولی با پاداش‌های در نظر گرفته شده متفاوت، وارد سیستم نموده و به اندازه‌گیری پارامترهایی نظیر زمان پاسخ‌دهی به سوالات، دقت پاسخ‌های ارائه شده و ... پرداخت. بررسی و تحلیل این پارامترها می‌تواند به روشن تر شدن اثر آن عامل بیانجامد. به عنوان مثال آزمایش‌های انجام شده در [۴۶] بر روی شبکه جمعی MTurk نشان داده‌اند که اگر هزینه‌های انجام ریز وظایف پایین آورده شوند، تاثیر چندان در صحت نتایج مساله نهایی بوجود نمی‌آید، ولی زمان انجام کار را بسیار طولانی‌تر می‌کند.

**حل ریز مسائل الگوریتمی.** واگذاری مسائل به شبکه انسانی هزینه‌بر می‌باشد، در نتیجه نیاز است که تا حد امکان نسبت به کمینه‌سازی تعداد مسائل واگذار شده به شبکه انسانی اقدام کنیم. یکی از راه‌ها برای انجام این امر، ماشینی کردن کارها تا حد امکان می‌باشد، یعنی شکست مساله به ریز مسائل، به قسمی که ریز مسائل تولیدی تا حد ممکن الگوریتمی باشند (چگونگی انجام این کار از مسائل پیش روی بخش «تولید ریز مسائل» می‌باشد). ریز مسائل الگوریتمی می‌توانند توسط ماشین حل شوند، در نتیجه هزینه چندان لازم نداشته و همچنین نیاز به اعتبارسنجی ندارند.

ارائه الگوریتم مناسب برای حل ریز مسائل الگوریتمی، کاملاً وابسته به مساله بوده و معمولاً این بخش جزو ساده‌ترین بخش‌های یک سیستم هوش جمعی می‌باشد. همچنین نحوه کارکرد در این بخش تأثیری بر بخش‌های دیگر سیستم ندارد.

**شبکه مورد استفاده.** در معماری یک سیستم متکی بر شبکه جمعی، ساختار شبکه (توپولوژی شبکه و نوع ارتباطات بین اعضا) نقش مهمی را ایفا می‌کند. اولین مساله‌ای هم که در این زمینه مطرح می‌شود، این است که آیا لازم است برای این مساله یک شبکه مختص به خود آن مساله ایجاد شود، یا اینکه استفاده از شبکه‌های جمعی یا اجتماعی موجود برای این امر می‌تواند کارا باشد؟

ایجاد یک شبکه جدید بسیار هزینه‌بر و زمان‌بر بوده، ولی از طرف دیگر، می‌تواند تطابق بسیار بیشتری بر نیازمندی‌های مساله مورد نظر داشته باشد. به همین دلیل، در حال حاضر اکثر پژوهش‌هایی که در این حوزه انجام می‌شوند، آزمایش‌های خود را بر روی شبکه‌های موجود و بخصوص شبکه MTurk<sup>۳</sup> آمازون انجام می‌دهند. این شبکه دارای بیش از یکصد هزار عضو بوده و مسائل بسیار زیادی به آن سپرده می‌شوند. همچنین این شبکه با استفاده از جعبه ابزاری که در اختیار برنامه‌نویسان و توسعه دهندگان قرار می‌دهد، امکان تعامل سیستم‌های دیگر را با خود امکان‌پذیر می‌سازد.

دلایل جذابیت MTurk برای محققان را از دید مقاله [۴۹] می‌توان بصورت زیر برشمرد: (۱) حجم بالای کاربران موجود برای محول کردن ریز مسائل (۲) واگرایی بالای کاربران از لحاظ دانش، فرهنگ، توانمندی، زبان، مکان، وضعیت اقتصادی اجتماعی و (۳۰۰۰) هزینه پایین به کارگیری کاربران این شبکه و (۴) چرخه سریع تئوری و تجربه (آزمایش). این مقاله همچنین به بیان فعالیت‌های تحقیقاتی متعددی می‌پردازد که برای انجام مسائل گسترده و متنوعی از MTurk استفاده کرده، و نتایجی هم سطح نتایج بدست آمده با سایر روش‌ها و حتی بهتر بدست آورده‌اند؛ در حالی که هزینه کمتری را نسبت به روش‌های جایگزین پرداخت کرده‌اند.

شبکه‌های دیگری نیز وجود دارند که برای این منظور مورد استفاده قرار گرفته‌اند. نمونه‌ای از آنها شبکه Tasken<sup>۴</sup> (معروف‌ترین وب‌سایت از نوع Witkey است که در چین راه‌اندازی شده‌اند) می‌باشد. این شبکه یک سایت اشتراک دانش بوده و در آن، بازدید کنندگان سوالات خود را مطرح کرده و هزینه‌ای که بابت دریافت بهترین پاسخ حاضرند بپردازند، را مشخص می‌کنند. کاربران شبکه برای دریافت پاداش مشخص شده با هم به رقابت پرداخته و پاسخ‌های خود را با بهترین کیفیت ارائه می‌نمایند. در نهایت درخواست کننده سوال، پاسخ مطلوب خود را مشخص کرده و هزینه مربوطه، به ارائه دهنده آن پاسخ تعلق می‌گیرد. مسائل ارجاعی به این شبکه از طراحی لوگو و توسعه وب‌سایت تا نوشتن طرح کسب‌وکار و از چند دهم دلار تا چند صد دلار متفاوتند. سایر مشخصات این شبکه عبارتند از: ۱.۷ میلیون کاربر (در مقایسه با جمعیت حدود یک‌صد هزار نفری MTurk)، تعداد وظایف ۳۱۰۰ (در مقایسه با چند صد هزار وظیفه MTurk)، تعداد پاسخ‌های ارائه شده ۵۴۳۰۰۰ تا ۲ سال، مشارکت حدود ۱۱ درصدی کاربران حداقلی در یکی از مسائل، دریافت پاداش توسط حدود ۳.۲ درصد کاربران مشارکت کننده و رقابت ۱۸۴ نفری کاربران در هر مساله [۵۰].

شبکه Tasken یک شبکه رقابتی بوده ولی MTurk یک شبکه غیر رقابتی می‌باشد. تحلیل اینکه رقابتی یا غیر رقابتی بودن یک شبکه چه تاثیری در خروجی نهایی می‌تواند داشته باشد، مبحث جالبی است که کمتر به آن پرداخته شده است. به عنوان مثال همانگونه که قبلاً نیز ذکر شد، نشان داده شده است [۴۶] که در MTurk تغییر هزینه تخصیص یافته برای یک مساله، کیفیت انجام آن را تحت تأثیر قرار نمی‌دهد، ولی در زمان انجام آن مؤثر است، در حالی که در [۵۰] نشان داده شده است که در Tasken افزایش هزینه، تأثیر مستقیم و زیادی در کیفیت انجام مساله دارد. به نظر می‌رسد دلیل اصلی این تفاوت در رقابتی بودن یا نبودن شبکه باشد (البته این نتیجه‌گیری نیازمند بررسی‌های بیشتر است). همچنین آزمایش‌های انجام شده بر روی شبکه Tasken و رفتارسنجی کاربران آن حول عامل رقابتی بودن شبکه نشان می‌دهند [۵۰] که بدلیل رقابت‌های فشرده، اکثر کاربران بعد از ارسال تعدادی محدودی پاسخ (حضور در تعداد محدودی رقابت) انگیزه‌های خود را از دست داده و غیر فعال می‌شوند. در نتیجه یکی از مشکلات اساسی این شبکه این است که فعال ماندن یک فرد در شبکه، کاملاً به این بستگی دارد که در طی تلاش‌های قبلی‌اش چند بار و هر بار بعد از چند تلاش، موفق به برنده شدن در یک رقابت شده است. همچنین رقابتی بودن شبکه باعث شده است که هسته بسیار محدودی از کاربران توانمند شکل بگیرد، که قادرند طوری برنامه‌ریزی کنند که هم در مسائل بیشتری برنده شوند و هم پاداش اکتسابی خود را بیشینه کنند (حدود یک پنجم جواب‌های برنده سایت از طرف این هسته ارائه می‌شوند). و نهایتاً اینکه، به این دلیل که کاربران، مطمئن به دریافت پاداش در مقابل کارشان نیستند، در مورد کاری که زمان زیادی نیاز داشته باشد، ریسک نمی‌کنند. در نتیجه در این شبکه، مسائل باید دارای پیچیدگی کمی بوده و تلاش چندان زیادی لازم نداشته باشند. جنبه مثبت رقابتی بودن نیز در این است که کاربران سعی می‌کنند وظایفی را برای پاسخ‌گویی انتخاب کنند که در ارائه پاسخ به آن‌ها، دارای مزیت رقابتی بوده و امکان احتمال برنده شدن آن‌ها بیشتر باشد. این امر باعث می‌شود که نیازی به هیچ‌گونه تخصیص وظیفه هوشمندی وجود نداشته باشد.

لازم به ذکر است که یک شبکه می‌تواند نیمه رقابتی نیز باشد. به عنوان نمونه‌ای از شبکه‌های نیمه رقابتی، می‌توان به شبکه بکار گرفته شده توسط MIT Media Lab در رقابت بالن‌های قرمز دارپا اشاره نمود، که در آن تمامی افرادی که در روند حل یک مساله شرکت داشته‌اند، پاداش دریافت کرده‌اند [۳].

علاوه بر نوع و جنس ریز مسائل هدف و رقابتی بودن یا نبودن شبکه، عوامل زیاد دیگری نیز وجود دارند که باید در طراحی شبکه مورد توجه قرار گیرند. این عوامل، و در نتیجه ساختار طراحی شده برای شبکه، بر نحوه تخصیص مسائل به اعضای شبکه، نحوه اعتبارسنجی پاسخ‌های ارائه شده برای مسائل و سایر استراتژی‌های بکار گرفته شده توسط سیستم تأثیر می‌گذارند. در ادامه، به بررسی چند مورد از این عوامل می‌پردازیم.

تعداد اجزای شبکه (شبکه یکپارچه یا شبکه با چند جزء مستقل): چند جزئی بودن شبکه می‌تواند استفاده‌های مفیدی داشته باشد، به عنوان مثال در یک شبکه دو جزئی، اعضای یکی از اجزا می‌توانند برای حل مساله مورد استفاده قرار گیرند و اعضای شبکه دیگر برای اعتبارسنجی. به عنوان مثال دیگری از کاربرد شبکه‌های چند جزئی می‌توان به چگونگی جلوگیری از کهنه شدن مسائل در سیستم‌های جمعی اشاره نمود. رفتارسنجی کاربران شبکه‌هایی که در آنها تمامی

<sup>۳</sup> Amazon Mechanical Turk, <http://www.mturk.com/>

<sup>۴</sup> <http://tasken.com>

ریز مسائل برای انتخاب توسط اعضا، در اختیار آنها قرار داده می‌شود، نشان می‌دهد که یکی از مهم‌ترین معیارهای کاربران برای انتخاب مساله، میزان تازگی مساله<sup>۵</sup> می‌باشد [۴۹]. در یک سیستم جمعی می‌توان دو شبکه مستقل از هم داشت، تا مسائلی که در یکی از شبکه‌ها کهنه می‌شوند، به شبکه دیگر تزریق شوند (و بالعکس)، تا مسائل در شبکه جدید دارای تازگی بوده و استقبال به آن‌ها بیشتر شود.

نوع ارتباطات بین اعضا: ارتباطات مختلفی می‌تواند بین اعضای یک شبکه جمعی وجود داشته باشد. به عنوان نمونه می‌توان ماتریس مجاورتی شبکه را به این صورت تعریف کرد که در ابتدا تمامی یال‌های شبکه دارای وزن صفر باشند. وقتی یک مساله یکسان برای حل به چند نفر سپرده می‌شود، در صورتی که پاسخ آنها مشابه بود، وزن لینک بین دو به دوی آنها را در گراف مجاورتی شبکه یک واحد افزایش و در غیر این صورت، یک واحد کاهش دهیم. در این صورت رابطه «تشابه نظرات» در شبکه ایجاد می‌شود. استفاده از اطلاعات این رابطه در شبکه می‌تواند در استراتژی‌های تخصیص وظیفه یا اعتبارسنجی پاسخ‌های کاربران بکار گرفته شود. تحلیل ویژگی‌های آماری انجمن‌های<sup>۶</sup> شکل گرفته بر اساس این یال‌ها نیز، می‌تواند اطلاعات مفیدی را از شبکه بدست دهد. همچنین به نظر می‌رسد که میزان تبادل اطلاعات در بین اعضا در یک سیستم هوش جمعی، نقش موثری در کارایی این شبکه داشته باشد. هنوز پژوهش خاصی در این زمینه‌ها انجام نشده است.

زیرساخت مورد استفاده: عامل دیگری که می‌تواند مطرح شود، زیرساخت‌های مورد استفاده شبکه می‌باشد. به عنوان مثال آیا شبکه از زیر ساخت وب بهره می‌گیرد یا یک شبکه بر پایه ارتباطات موبایلی می‌باشد؟ این امر به این دلیل می‌تواند اهمیت داشته باشد که در یک شبکه بر پایه ارتباطات موبایلی می‌توان از اطلاعات مکانی و سنسورهای تلفن‌های همراه کاربران برای حل مسائل مبتنی بر مکان<sup>۷</sup> استفاده نمود، در حالی که استفاده از شبکه‌های تحت وب برای این امر چندان منطقی نیست. شبکه txteagle<sup>۸</sup> نمونه‌ای از شبکه‌های بر پایه ارتباطات موبایلی می‌باشد.

عوامل مختلف دیگری نیز در این زمینه می‌توانند ذکر شوند. شناسایی سایر عوامل یا سنجش تاثیر عوامل مختلف بر سایر بخش‌ها، می‌تواند با رفتارسنجی کاربران شبکه‌های مختلف، بوسیله آزمایش‌های کنترل شده انجام پذیرد.

**تخصیص ریز مسائل.** یک سوال پایه‌ای مهم در رابطه با انجام مسائل توسط کاربران این است که آیا بهتر است تمامی مسائل مورد نیاز برای پاسخ‌گویی در اختیار تمامی کاربران قرار گیرد، تا خود از بین آنها مسائلی را برای پاسخ دادن انتخاب کنند، یا بهتر این است که سیستم، با توجه به کارکرد و سابقه افراد، مسائل را به آنها تخصیص دهد. یا اینکه ترکیبی از این دو استراتژی بکار گرفته شود.

از بین شبکه‌های موجود، شبکه MTurk و Tasken از رویکرد اول استفاده می‌کنند، ولی در txteagle تخصیص مسائل به کاربران توسط سیستم صورت می‌پذیرد. از این دید txteagle یک سیستم فعال<sup>۹</sup> بوده، ولی MTurk و Tasken سیستم‌های منفعل<sup>۱۰</sup> می‌باشند. در رویکرد اول، با وجود اینکه تخصیص مساله به کاربر به صورت صریح وجود ندارد، ولی بصورت ضمنی می‌تواند از طریق بررسی و تحلیل استراتژی‌های سود بهینه اتخاذ شده توسط کاربران، انجام می‌پذیرد. به این صورت که با تحلیل رفتار کاربران مختلف در انتخاب مسائل برای انجام، می‌توان استراتژی انتخاب مسائل گروه‌های مختلف کاربران را شناسایی نموده و مسائل را طوری طراحی نمود، که عملاً کاربران مورد نظر، نسبت به انتخاب این مسائل از خود اشتیاق نشان دهند. در اینجا طراحی مساله شامل نحوه ورود مسائل به مخزن<sup>۱۱</sup> مسائل فعال، زمان در نظر گرفته شده برای هر مساله، پاداش اختصاص داده شده به آن و هر پارامتر دیگری می‌باشد، که در تحلیل رفتاری کاربران در انتخاب مساله مهم تشخیص داده شده است.

در رویکرد دوم، این واقعیت مورد استفاده قرار می‌گیرد، که هر عضو یک سیستم هوش جمعی، دارای سابقه‌ای در آن سیستم است. این سابقه می‌تواند توسط سیستم مدل شده و مورد استفاده قرار گیرد. به نظر می‌رسد که در حالت‌هایی که ریز مسائل موجود، دارای وابستگی یا اولویت‌بندی باشند، تخصیص وظیفه صریح، اهمیت بیشتری پیدا کند.

لازم به ذکر است که استراتژی‌های انتخاب مساله برای تزریق به شبکه و اعتبارسنجی پاسخ‌های ارائه شده توسط کاربران، می‌تواند مستقیماً از استراتژی بکار گرفته شده در این بخش تاثیر پذیرد. از آن جایی که تحقیقات چندانی در این بخش انجام نشده است، مسائل تحقیقاتی متنوعی می‌تواند در این اینجا مطرح شود.

**مدل‌سازی اعضا.** دلایل زیادی مبنی بر لزوم نگهداری سوابق اعضای سیستم و مدل‌سازی رفتار آنها وجود دارد. اولین دلیل می‌تواند شناسایی اعضای با کارایی پایین یا منبع تولید نویز و خطا در سیستم باشد. دومین دلیل آن می‌تواند استفاده از این اطلاعات در «تخصیص وظایف به کاربران» باشد که در بخش

<sup>۵</sup>Task recency<sup>۶</sup>Community<sup>۷</sup>Location based problems<sup>۸</sup><http://txteagle.com/><sup>۹</sup>Active<sup>۱۰</sup>Passive<sup>۱۱</sup>Pool

قبلی مورد بررسی قرار گرفت. سومین دلیل آن، همانطور که در بخش‌های بعدی خواهیم دید، استفاده از این اطلاعات در اعتبارسنجی پاسخ‌های ارائه شده توسط کاربران سیستم می‌باشد. آخرین دلیل نیز، استفاده از این اطلاعات در محاسبه احتمال درستی پاسخ نهایی یک مساله می‌باشد.

طبقه‌بندی کاربران و مدل‌سازی مهارت‌های آن‌ها در کنار طبقه‌بندی مسائل، می‌تواند تأثیر زیادی در نتیجه نهایی داشته باشد. به عنوان مثال در [۴۶] با وجود اینکه در تصحیح و ارتقای سطح کیفی متون عمومی نتایج بالایی اخذ شده است، ولی بدلیل عدم تخصصی بودن شبکه کاربران جمعی بکار گرفته شده (یا در اصل عدم تخصیص مسائل تخصصی به کاربران با تخصص مرتبط)، نتایج بدست آمده در متون تخصصی ضعیف بوده است (منطبق بر دلیل دوم: نحوه تخصیص وظیفه به کاربران).

مدل‌سازی اعضای شبکه می‌تواند از یک فراوانی نمرات کسب شده قبلی در هر طبقه از مسائل، تا مدل‌های بسیار پیچیده‌تری که ویژگی‌های رفتاری آنها را در مواجهه با مسائل مختلف نشان دهد، متفاوت باشد. این بخش یکی از اساسی‌ترین بخش‌های چارچوب می‌باشد، که باید هم‌زمان با سایر بخش‌های مرتبط مورد بررسی قرار گیرد. در مورد این بخش، در فصل بعد، بیشتر صحبت خواهیم کرد.

**اعتبارسنجی پاسخ‌ها.** از مهم‌ترین بخش‌ها در سیستم‌های شامل عامل‌های انسانی، بررسی خروجی آن عامل‌ها می‌باشد. روش‌هایی که برای این منظور به کار گرفته می‌شوند، بستگی به عوامل مختلفی دارند. پیچیدگی اصلی سیستم‌های محاسبات جمعی در این بخش می‌باشد. این بخش باید هم‌زمان با سایر بخش‌های مرتبط مورد بررسی قرار گیرد. در مورد این بخش، در فصل بعد، به طور مفصل صحبت خواهیم کرد.

**تجمیع پاسخ‌ها.** این بخش وابسته به مساله بوده و می‌تواند از یک ذخیره‌سازی ساده در مثال بازیابی تصویر ارائه شده تا روش‌هایی نظیر Bayesian Truth Serum [۱۰] در پیش‌بینی جمعی متفاوت باشد. بدلیل وابستگی بالای این بخش به جنس مساله، توضیحات بیشتری در مورد آن ارائه نخواهیم داد.

**بازه اطمینان پاسخ نهایی** هدف از این بخش این است که احتمال درستی پاسخ نهایی ارائه شده برای مساله اصلی را محاسبه نماید. این بخش کاملاً متأثر از نحوه اعتبارسنجی پاسخ‌های شبکه می‌باشد. در مورد این بخش همراه با مبحث اعتبارسنجی پاسخ‌ها، در فصل بطور مفصل صحبت خواهیم کرد.

در این فصل، بخش‌های مختلف یک سیستم جمعی را مورد شناسایی قرار دادیم. در ادامه این پروژه به بررسی بخش‌های «تولید ریز مسائل و تخصیص آنها به اعضای شبکه، مدل‌سازی اعضا، اعتبارسنجی پاسخ‌ها و بازه اطمینان پاسخ نهایی» در کنار هم می‌پردازیم؛ تا نحوه بکارگیری از آنها و تأثیرات متقابلشان را بر روی هم، در راستای هدف «؟؟؟» مورد بررسی قرار دهیم.

## فصل ۳

# مرور ادبیات موضوع

در این فصل به بررسی ادبیات موضوع در حوزه استراتژی حل مساله با استفاده از جمع‌سپاری می‌پردازیم. استراتژی حل مساله شامل چند بخش می‌باشد. اول اینکه با توجه به ماهیت مساله و ویژگی‌های کاربران موجود، چگونه می‌توان از کاربران شبکه برای حل یک مساله بهره برد؟ دوم اینکه با داشتن تعداد زیادی سوال، چگونه و در چه زمان‌هایی، کدام سوالات را به منبع سوالات فعال شبکه تزریق کنیم؛ یا در صورتی که مسائل به کاربران به صورت خودکار تخصیص می‌یابند، کدام مساله به کدام کاربر تخصیص داده شود. و نهایتاً اینکه چگونه نسبت به اعتبارسنجی پاسخ‌های دریافتی از کاربران و محاسبه احتمال درستی پاسخ نهایی بدست آمده اقدام نمائیم. در ادامه به بررسی هر کدام از این موارد می‌پردازیم.

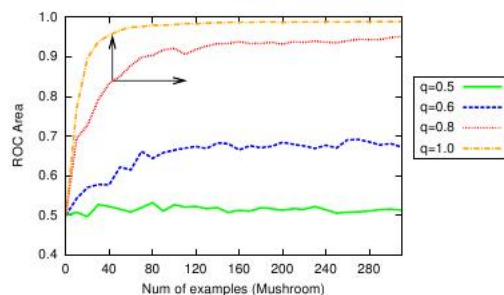
### ۳.۱ نحوه استفاده

شبکه‌های جمعی به شیوه‌های مختلفی می‌تواند در حل مسائل بکار گرفته شوند. ابتدایی‌ترین شیوه این است که هر مساله مستقیماً به شبکه داده شده و جواب آن از شبکه استنتاج شود. در حالت‌های دیگر، پاسخ‌های دریافت شده برای برخی از سوالات می‌توانند برای آموزش یک مدل بکار گرفته شوند، به قسمی که مدل قادر باشد، به تنهایی و بدون نیاز به شبکه جمعی به دیگر سوالات کاربران به طور مستقل پاسخ دهد. در این رویکرد می‌توان از دو استراتژی طراحی مدل ثابت و یادگیری فعال بهره برد که در ادامه به بررسی آنها می‌پردازیم.

**طراحی مدل ثابت.** در این روش، هدف، بدست آوردن پاسخ برای یک مجموعه آموزشی بدون پاسخ جهت ساخت یک مدل می‌باشد. در این رویکرد به جای اینکه تمامی داده‌ها برای حل به شبکه سپرده شوند، برای کاهش هزینه‌ها، فقط تعدادی از داده‌ها، تحت عنوان مجموعه آموزشی، به این شبکه سپرده می‌شوند. در مرحله بعد با توجه به پاسخ‌های بدست آمده برای داده، یک مدل آموزش داده شده و بقیه مسائل برای حل به جای شبکه، به این مدل داده می‌شوند. اگر سوالات داده‌های آموزشی نمونه خوبی از کل جمعیت سوالات باشند، استفاده از مدل حاصل به تنهایی، پاسخ‌های خوبی را بر می‌گرداند، علاوه بر اینکه صرفه‌جویی بالایی نیز در هزینه‌ها انجام می‌شود.

بدیهی است که کیفیت مدل بدست آمده به تعداد داده‌های آموزشی و همچنین کیفیت کار کاربران شبکه (احتمال پاسخ‌گویی درست کاربران به مسائل) بستگی دارد. این تاثیر برای یک مساله کلاسه‌بندی در شکل ۳.۱ برای حالتی که احتمال پاسخ‌گویی درست به سوالات برای تمامی کاربران مساوی هم و برابر  $q$  در نظر گرفته شود، به ازای  $q$ های مختلف، نشان داده شده است. نتیجه مهمی که می‌توان از این شکل گرفت این است که به شرط  $q$  مناسب، هرچه تعداد داده‌های آموزشی بالاتر رود، کیفیت نهایی کلاسه‌بند بالاتر می‌رود؛ و دیگر اینکه هر چه  $q$  بالاتر رود نتیجه نهایی کلاسه‌بند بهتر می‌شود. پس اگر برای یک کلاسه‌بند مشخص بخواهیم بهبود کیفیت ایجاد کنیم، یکی از کارهای زیر را باید انجام دهیم: یا برای تعداد نمونه‌های بیشتری درخواست پاسخ کنیم که تعداد نمونه‌های آموزشی بالاتر رود (پیکان افقی در شکل ۳.۱). یا کیفیت پاسخ‌گویی کاربران شبکه را افزایش دهیم (پیکان عمودی در شکل ۳.۱). در بسیاری از مسائل، این امر می‌تواند از طریق سپردن هر مساله به تعدادی بیشتری از کاربران حاصل شود. یا اینکه می‌توان ترکیبی از دو عمل فوق را انجام دهیم [۵۷]. تحلیل اینکه در چه شرایطی، استفاده از کدام استراتژی یا چه ترکیبی از این دو استراتژی، چه بهبودی در کیفیت کلاسه‌بند ایجاد می‌کند، مبحثی است که فعالیت خاصی در مورد آن انجام نشده است مساله باز.

**یادگیری فعال.** ایده اصلی یادگیری فعال این است که یک مدل موجود (مثلاً با استفاده از رویکرد قبلی)، در بسیاری از مواقع می‌تواند با نمونه‌های آموزشی



شکل ۳.۱: اثر تعداد داده‌های آموزشی و احتمال درستی پاسخ‌های ارائه شده توسط کاربران بر کیفیت کلاسه‌بند [۵۷].

جدید تقویت شود. در یادگیری فعال با مشاهده هر نمونه ورودی جدید، یا آن نمونه برای حل به مدل داده می‌شود، یا اینکه برای تقویت مدل فعلی بکار گرفته شده و جهت دریافت پاسخ برای آن به شبکه جمعی سپرده می‌شود. مساله اصلی در این رهیافت استراتژی بکار گرفته شده برای تصمیم‌گیری فوق می‌باشد.

به عنوان نمونه‌ای از این کاربرد می‌توان به مقاله [۵۵] اشاره نمود که به کلاسه‌بندی برخط تویت‌ها<sup>۱</sup> با استفاده از این رویکرد می‌پردازد. ویژگی‌های مورد استفاده برای کلاسه‌بندی تویت‌ها، کلمات بکار رفته در متن تویت می‌باشند. اگر تویت حاوی کلماتی باشد که کلاسه‌بند را قادر سازند که به صورت کارایی کلاسه‌بندی را انجام دهد، این عمل انجام می‌شود. در غیر این صورت این تویت به عنوان یک داده آموزشی برای تقویت کلاسه‌بند بکار می‌رود. برای لیبل‌گذاری این داده جهت استفاده به عنوان داده آموزشی، از شبکه MTurk و استراتژی رأی اکثریت استفاده شده است (حالت مساله کلاسه‌بندی دو کلاسه با کیفیت کاربران ثابت، که در ادامه توضیح داده خواهد شد).

## ۳.۲ تزریق و تخصیص مسائل

تزریق و تخصیص مسائل در استراتژی‌های استفاده مختلف به صورت‌های مختلفی مطرح می‌شوند. در استفاده مستقیم هدف این است که مسائل طوری به سیستم تزریق شوند یا به کاربران آن اختصاص یابند که احتمال درستی پاسخ نهایی حاصله بیشینه باشد. مسائل مطرح در این دیدگاه عبارتند از اینکه چه زمانی و با چه ترتیبی مسائل را به سیستم تزریق کنیم؟ هر مساله باید به چند کاربر داده شود و .... در استراتژی طراحی مدل ثابت، مساله دیگری نیز علاوه بر این موارد مطرح می‌باشد که عبارت است از اینکه هزینه‌بندی باید به چه صورت انجام شود، تا کیفیت مدل استخراجی بهینه شود. و در استراتژی یادگیری فعال، علاوه بر موارد فوق، هدف دیگری نیز مطرح است که عبارت است از تصمیم‌گیری برای هر مساله ورودی، که آیا لازم است که آن مساله برای تقویت کلاسه‌بند استفاده شده و به سیستم تزریق شود یا خیر؟ در ادامه به بررسی موارد فوق می‌پردازیم.

### ۱.۳.۲ استفاده مستقیم

برای سیاست‌گذاری‌های مربوطه در این خصوص، می‌توان از رفتارسنجی‌های کاربران این شبکه‌ها استفاده نمود. به عنوان مثال، در بررسی رفتاری کاربران MTurk مشخص شده است که آن‌ها از معیارهای تازه‌گی مسائل در سیستم (زمان اضافه شدن آنها به شبکه) و تعداد مسائل هم‌جنس در شبکه بهره می‌گیرند تا مساله‌ای را برای پاسخ‌گویی انتخاب نمایند. همچنین بر خلاف تصور، هزینه اختصاص یافته به مساله نه تنها به عنوان یک معیار مثبت مطرح نیست، بلکه دارای نقش منفی نیز می‌باشد. دلیل آن نیز این است که کاربر می‌داند که سوالی که هزینه بالایی برای پاسخ‌گویی به آن در نظر گرفته شده است، کار زیادی برده و ارزش انتخاب ندارد. البته اگر هزینه مسائل از حد معینی هم کمتر شود، کاربران از آن چشم‌پوشی می‌کنند. نقش موثر هزینه در کاهش زمان انتخاب و پاسخ‌گویی به سوالات از طرف کاربران می‌باشد [۵۹].

در نتیجه می‌توان از مکاشفه‌هایی که از مطالعات رفتارسنجی بدست می‌آید، در طراحی کاراتر سوالات و چگونگی تزریق آنها به شبکه بهره برد. مثلاً بهتر است مسائل مختلف از یک جنس در چند مرحله به سیستم تزریق شوند، تا در هر لحظه تعدادی از آنها دارای تازه‌گی بوده و مورد استقبال واقع شوند. یا تا حد امکان مسائل را با یک ساختار یکسان طراحی نمود، تا نیاز به مرحله آموزش برای حل سوالات مختلف تا حد ممکن کاهش یابد، همچنین با توجه به تجارب بدست آمده در مورد هزینه‌ها، استراتژی مناسب می‌تواند شروع با هزینه وظیفه حداقلی، و افزایش آن در صورت عدم استقبال باشد و ....

<sup>۱</sup>Twittes

## ۲.۳.۲ مدل ثابت و هزینه بندی

درخواست هر پاسخ از سیستم، هزینه مشخصی را در بر دارد. از طرفی بودجه در نظر گرفته شده برای تشکیل مجموعه آموزشی به منظور طراحی یک مدل یا کلاسه بندی، محدود و مشخص می باشد. سوالی که در اینجا مطرح می شود این است که آیا بهتر است برای نمونه های بیشتری هر کدام پاسخ های کمتری داشته باشیم، یا برای نمونه های کمتری، هر کدام پاسخ های بیشتری؟ پاسخ های بیشتر منجر به احتمال درستی پاسخ نهایی بالاتر برای مساله مورد نظر می شوند. از طرفی ممکن است افزایش احتمال درستی پاسخ نهایی که با اضافه کردن یک پاسخ بیشتر حاصل می شود، خیلی از کمتر از اضافه کردن یک نمونه آموزشی جدید به سیستم باشد. پس در اینجا با یک مساله بهینه سازی سر و کار خواهیم داشت. مساله آنجا پیچیده تر خواهد شد که تعداد پاسخ های درخواستی برای تمامی مسائل برابر نباشد. در این صورت باید به سوال دیگری هم پاسخ داده شود و آن این است که به کدام مسائل و با چه توجهی و به چه تعدادی باید پاسخ های بیشتری درخواست کنیم؟ این تصمیم گیری ها به چه عواملی بستگی دارد؟ و با توجه به آن عوامل به چه نحو می تواند در مورد آن تصمیم گیری شود؟ مساله باز

استراتژی های مختلف برای این منظور می تواند مطرح باشد که در ادامه به بررسی آنها می پردازیم. بدیهی ترین استراتژی، درخواست تنها یک پاسخ به ازای هر سوال می باشد. در این حالت مساله تصمیم گیری خود به خود حذف خواهد شد. استراتژی دیگر، درخواست تعداد ثابتی پاسخ برای هر سوال می باشد. در این حالت به ترتیب و به صورت گردشی<sup>۲</sup> برای تمامی سوالات، درخواست پاسخ می شود، تا زمانی که بودجه مورد نظر پایان پذیرد. در حالی که نویز و اشتباه کاربران بالا باشد، یعنی احتمال درستی پاسخ توسط کاربران ( $p$ ) پایین باشد (بدیهی است که برای هر کاربر باید  $p > 0.5$  باشد)، این استراتژی به شرط داشتن تعداد داده های آموزشی کافی، جواب های خوبی را بدست می دهد. ولی اگر تعداد پاسخ ها به ازای هر نمونه بیشتر از حد معینی شود، کیفیت پاسخ های بدست آمده دیگر بالاتر نمی رود و از آن به بعد تعداد داده های آموزشی است که اهمیت پیدا می کند. برای  $p$  های بزرگ نقش تعداد داده ها از همان ابتدا بسیار اساسی تر است، چون در این حالت، boosting تأثیر چندانی در بالا بردن دقت پاسخ گویی ندارد، ولی در عوض، تعداد نمونه های آموزشی نقش موثری را در این زمینه ایفا می کند [۵۷].

آخرین استراتژی نیز درخواست تعداد پاسخ های متغیر برای سوال های متفاوت می باشد. در این استراتژی لزوماً تعداد پاسخ های دریافت شده برای هر سوال به عنوان معیاری برای درخواست پاسخ بعدی مورد استفاده قرار نمی گیرد (نظیر استراتژی قبلی). بلکه هر معیار دیگری می تواند برای این منظور به کار گرفته شود.

یک معیار مناسب برای این منظور می تواند میزان ناخالصی (همگنی) مجموعه پاسخ های موجود فعلی باشد. برای سنجش میزان خلوص یک مجموعه، می توان از دو معیار سود جست: اول اینکه آیا رأی اکثریت به حد آستانه مورد پذیرش (مثلاً ۵۰ درصد) رسیده است؟ و دوم اینکه آنتروپی آرای جمع شده چقدر است؟

آزمایش ها [۵۷] نشان می دهند که بکارگیری آنتروپی پاسخ های موجود فعلی، نتایج خوبی را به دنبال نخواهد داشت. دلیل این امر، نویزی بودن پاسخ های کاربران می باشد. زیرا هم عقیده بودن تعداد محدودی از کاربران در مورد پاسخ یک سوال، دال بر درست بودن آن پاسخ نیست. این مشکل جایی خود را بیشتر می نمایاند که کیفیت کار کاربران پایین باشد [۵۷] یا تعداد پاسخ های موجود، کم باشد [۵۶].

یکی از معیارهای دیگر برای این منظور می تواند محاسبه عدم قطعیت جواب بدست آمده از برآیند پاسخ های جمع شده فعلی باشد. آزمایش ها [۵۷] نشان می دهند که استفاده از عدم قطعیت در حالت نویزی (کیفیت پایین کار کاربران) نتیجه بهتری را نسبت به استفاده از آنتروپی بر می گرداند، ولی هر چه این کیفیت بالاتر رود، بهبود نتیجه با دخیل کردن عدم قطعیت کاهش می یابد.

در ادامه به بررسی نحوه محاسبه عدم قطعیت در دو حالت مختلف می پردازیم. توضیح اینکه تمامی روش هایی که در ادامه آمده اند بر اساس استفاده از تکنیک رأی اکثریت طراحی شده اند. بدیهی است که در تمامی این روش ها، هر چه عدم قطعیت جواب نهایی استخراج شده برای یک سوال بالاتر باشد، آن سوال بیشتر نیازمند دریافت پاسخی جدید می باشد.

### تکنیک رأی اکثریت بدون استفاده از کیفیت کاربران

در [۵۶] عدم قطعیت برای یک مساله کلاسه بندی دو کلاسه، از طریق احتمال برابر نبودن رأی اکثریت با پاسخ درست محاسبه شده است. همچنین برای محاسبه این احتمال هیچ استفاده ای از اطلاعات کاربران شبکه (احتمال درستی پاسخ های ارائه شده از طرف کاربران) نشده است. در ادامه به بررسی نحوه محاسبه این احتمال می پردازیم.

داده نمونه  $x$ : با کلاس واقعی  $y$  را در نظر بگیرید. همچنین فرض کنید تعدادی رأی از کاربران جمع شده است که  $L_1$  تای آن ها، نظر به تعلق نمونه به کلاس + و  $L_2$  آن ها، نظر به تعلق نمونه به کلاس - دارند. و فرض کنید کلاس مورد نظر اکثریت  $\hat{y}$  باشد.

<sup>۲</sup>Round robin

توزیع احتمال پیشین  $y$  یکنواخت در نظر گرفته شده است. با استناد به [۷۸] احتمال پسین  $y$  توزیعی بتا با پارامترهای  $(L_1 + 1, L_2 + 1)$   $\beta$  (با فرض  $L_1 > L_2$ ) می‌باشد. برای انتخاب احتمال پسین از توزیع بدست آمده در این مقاله ظاهراً از میانه<sup>۲</sup> توزیع استفاده شده است (توضیح دقیقی در این مورد داده نشده است. فقط حد آستانه تصمیم‌گیری برابر ۰.۵ ذکر شده است که در این مثال دلالت بر انتخاب میانه دارد. بدیهی است که دلیلی نیز برای ترجیح میانه بر میانگین یا مد ارائه نشده است). اگر میانه از ۰.۵ کمتر باشد، داده متعلق به کلاس + و در غیر این صورت متعلق به کلاس - در نظر گرفته شده است. میزان عدم قطعیت را نیز برابر مقدار کمتر بین دو مساحت زیر نمودار از صفر تا نیم و از نیم تا یک در نظر گرفته است (در اینجا نیز هیچ توضیحی ارائه نشده است که چرا این مرز برابر ۰.۵ در نظر گرفته شده است و نه برابر میانه). برای محاسبه مساحت زیر نمودار از توزیع تجمیعی<sup>۳</sup> توزیع بتا استفاده شده است. توزیع تجمیعی بتا برابر است با تابع بتای ناکامل تنظیم شده<sup>۴</sup> که با  $I_x$  نمایش داده می‌شود. در نتیجه میزان عدم قطعیت در این روش برابر است با:

$$S_{LU} = \min\{I_{0.5}(L_1 + 1, L_2 + 1), 1 - I_{0.5}(L_1 + 1, L_2 + 1)\}$$

آزمایش‌های انجام شده در این مقاله نشان می‌دهند که معیار انتخاب «عدم قطعیت پاسخ» جواب‌های بهتری را به نسبت معیار انتخاب «کمترین تعداد پاسخ‌ها» بدست می‌دهد.

### تکنیک رای اکثریت و کاربران با کیفیت یکسان

این روش در [۵۷] ارائه شده و همانند روش قبلی عدم قطعیت پاسخ برای یک مساله کلاسه‌بندی دو کلاسه، از طریق احتمال برابر نبودن رأی اکثریت با پاسخ درست محاسبه شده است. ولی بر خلاف روش قبلی از احتمال درستی پاسخ ارائه شده توسط کاربران استفاده شده است.

اگر کیفیت کار تمام کاربران را ثابت و برابر  $p$  در نظر بگیریم، می‌توان  $p$  را از روی فعالیت فعلی کاربران محاسبه نمود. با توجه به توضیحات ارائه شده در روش قبلی (و با فرض  $L_1 > L_2$ ) می‌توان توزیع احتمال  $p$  را بصورت  $P(+|L_1 + 1, L_2 + 1)$  تخمین زد.

همچنین می‌توان  $P(+|L_1, L_2)$  و  $P(-|L_1, L_2)$  را بصورت زیر نوشت:

$$P(+|L_1, L_2) = \frac{P(L_1, L_2|+)P(+)}{P(L_1, L_2)} = p^{L_1} \cdot (1-p)^{L_2} \cdot \frac{P(+)}{P(L_1, L_2)}$$

$$P(-|L_1, L_2) = p^{L_2} \cdot (1-p)^{L_1} \cdot \frac{P(-)}{P(L_1, L_2)}$$

در این صورت با انتگرال بر روی تمامی مقادیر  $p$  می‌توان احتمال  $P(+|L_1, L_2)$  را بصورت زیر نوشت:

$$P(+|L_1, L_2) = P(L_1, L_2|+) \cdot \frac{P(+)}{P(L_1, L_2)}$$

$$= \frac{P(+)}{P(L_1, L_2)} \int_0^1 p^{L_1} \cdot (1-p)^{L_2} \cdot P(p) dp$$

$$= \frac{P(+)}{P(L_1, L_2)} \int_0^1 p^{L_1} \cdot (1-p)^{L_2} \cdot \beta(p; L_1 + 1, L_2 + 1) dp$$

$$= \frac{P(+)}{P(L_1, L_2)} \int_0^1 p^{L_1} \cdot (1-p)^{L_2} \cdot \frac{\Gamma(L_1 + L_2 + 2)}{\Gamma(L_1 + 1)\Gamma(L_2 + 1)} \cdot p^{L_1} \cdot (1-p)^{L_2} dp$$

$$= \frac{P(+)}{P(L_1, L_2)} \cdot \frac{\Gamma(L_1 + L_2 + 2)}{\Gamma(L_1 + 1)\Gamma(L_2 + 1)} \int_0^1 p^{2L_1} \cdot (1-p)^{2L_2} dp$$

$$= \frac{P(+)}{P(L_1, L_2)} \cdot \frac{\Gamma(L_1 + L_2 + 2)}{\Gamma(L_1 + 1)\Gamma(L_2 + 1)} \cdot \frac{\Gamma(2L_1 + 1)\Gamma(2L_2 + 1)}{\Gamma(2L_1 + 2L_2 + 2)}$$
(۱.۳)

با نوشتن  $P(-|L_1, L_2)$  به همین صورت و توجه به اینکه مجموع این دو باید برابر ۱ باشد و همچنین اینکه  $L_1$  و  $L_2$  اعداد صحیح هستند و با توجه به

<sup>۲</sup> median

<sup>۳</sup> CDF

<sup>۴</sup> Regularized incomplete Beta function



تعریف تابع گاما می‌توان این احتمال را بصورت زیر ساده‌سازی نمود:

$$P(+|L_1, L_2) = \left( 1 + \frac{1 - P(+)}{P(+)} \cdot \frac{((L_1 + L_2)!)^2}{(2L_1)!(2L_2)!} \right)^{-1} \quad (۲.۳)$$

تنها مجهول رابطه فوق  $P(+)$  می‌باشد که برای محاسبه آن می‌توان از یک EM ساده بصورت زیر بهره برد:

۱. یک مقدار اولیه تصادفی برای  $P(+)$  در نظر بگیر (مثلاً ۰.۵).

۲. برای تمامی نمونه  $P(y_i = +|L_1, L_2)$  را با استفاده از احتمال پیشین کلاس مثبت فعلی محاسبه کن.

۳. با توجه به احتمال کلاس نمونه‌ها،  $P(+)$  جدید را با میانگین‌گیری از روی آن‌ها محاسبه کن.

۴. به مرحله ۲ برو.

و در نهایت میزان عدم قطعیت برابر خواهد بود با:

$$S_{LU} = \min\{P(+|L_1, L_2), P(-|L_1, L_2)\}$$

با توجه به استفاده از احتمال پیشین کلاس‌ها، این روش در مسایلی خوب جواب می‌دهد که توزیع کلاس‌ها نامتوازن باشد.

قبل از به پایان رسانیدن این بخش، به ذکر چند موضوع دیگر در رابطه با هزینه‌بندی می‌پردازیم که کمتر به آنها پرداخته شده است. این موضوعها عبارتند از: آیا فقط مسائل تکمیل نشده می‌توانند برای حل انتخاب شوند یا یک پاسخ مجدد می‌تواند برای یک مساله حل شده قبلی نیز درخواست گردد؟ یا اینکه اگر در هنگام انتخاب سوال برای درخواست جواب، چند سوال دارای مقدار یکسانی برای معیار مورد استفاده بودند، انتخاب از بین آنها به چه شکلی صورت پذیرد؟ یا اینکه انتخاب ما متعین باشد یا تصادفی، به این معنا که همیشه سوال با بالاترین مقدار برای معیار تصمیم‌گیری را انتخاب کنیم یا اینکه مثلاً مقادیر معیار را برای همه سوالات مشخص نمائیم و یکی از سوالات را به تصادفی با احتمال نسبت مقدار معیارش برگزینیم؟ یا ... فعالیت‌هایی نیز در این زمینه انجام شده است که به عنوان مثال می‌توان به آزمایش‌ها انجام شده در [۵۷] اشاره نمود که در آن نشان داده شده است که انتخاب تصادفی به نسبت مقدار معیار، نسبت به حالت متعین نتایج ضعیف‌تری را بر می‌گرداند. دلیل آن نیز این است که تعداد نمونه‌ها با عدم اطمینان پایین بالا می‌باشد و همین امر سبب می‌شود تا احتمال انتخاب نمونه‌های با عدم قطعیت بالاتر افت پیدا کند.

## ۳.۳.۲ یادگیری فعال

همانگونه که پیش‌تر نیز ذکر شد، یکی از روش‌های استفاده از شبکه‌های جمعی برای حل مسائل، استفاده از یادگیری فعال می‌باشد. سه تفاوت عمده بین یادگیری فعال در اینجا و روش‌های معمول یادگیری فعال وجود دارند: یکی اینکه در اینجا هزینه تشکیل مجموعه آموزشی، بدلیل اینکه توسط عوامل انسانی انجام می‌پذیرد، رایگان نیست. در نتیجه تعداد درخواست‌ها برای حل مسائل به عنوان داده آموزشی باید پایین باشد و همچنین شروع روش باید با تعداد بسیار محدودی داده آموزشی صورت پذیرد. دومین وجه تفاوت این است که داده‌های جدید بصورت جریان<sup>۶</sup> وارد می‌شوند و مثل روش‌های معمول، این امکان را نداریم که تعداد زیادی داده در اختیار داشته باشیم و از بین آن‌ها بخواهیم تعدادی را برای آموزش و تقویت مدل موجود انتخاب کنیم. داده‌ها بصورت جریان وارد شده و برای هر داده جدید، فقط با توجه به خود آن داده، باید تصمیم بگیریم که آیا آن را برای تقویت کلاسه‌بند و حل توسط عوامل انسانی انتخاب کنیم یا به عنوان یک داده تست برای ورود به کلاسه‌بند؟ تفاوت دیگری که این مساله با یادگیری فعال دارد این است که در یادگیری فعال پاسخ‌های نمونه‌های آموزشی دارای نویز نیست. هر سه عامل فوق خود را در مساله تصمیم‌گیری برای ورود به شبکه جمعی، یا به مدل نشان می‌دهند.

### ترکیب مدل‌ها

یک روش مناسب برای تصمیم‌گیری در مورد یک نمونه، این است که ابتدا آن نمونه به مدل طراحی شده داده شده و احتمال نادرستی پاسخ ارائه شده توسط مدل سنجیده شده و به عنوان عدم قطعیت سیستم برای پاسخ ارائه شده در نظر گرفته شود. در [۵۶] برای این منظور از یادگیری تعدادی ( $m$ ) مدل استفاده

<sup>۶</sup>stream

شده است. میزان عدم قطعیت در این حالت بصورت زیر سنجیده می‌شود (در این روش فرض می‌شود که مدل‌های موجود ضمن برگرداندن پاسخ سوال، احتمالی درستی آن را نیز بر می‌گرداند):

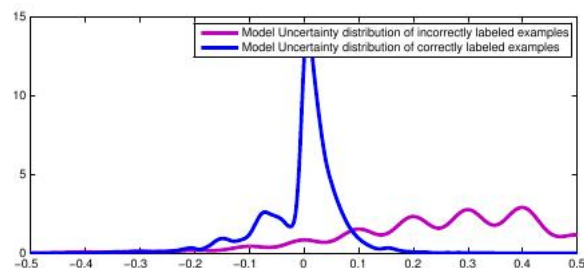
$$S_{MU} = \frac{1}{2} - \left| \frac{1}{m} \sum_{i=1}^m P(+|x, H_i) - \frac{1}{2} \right| \quad (۳.۳)$$

در حالتی که میزان عدم قطعیت بدست آمده فوق بالا باشد، می‌توان تعدادی پاسخ نیز برای سوال مورد، از شبکه جمعی نظر بدست آورده و با پاسخ مدل‌ها تلفیق نمود. عدم قطعیت مدل (رابطه فوق) با عدم قطعیت پاسخ‌ها (که در بخش‌های قبلی معرفی شد)، می‌تواند طبق رابطه زیر ترکیب شوند:

$$S_{LMU} = \sqrt{S_{MU} \cdot S_{LU}} \quad (۴.۳)$$

آزمایش‌های انجام شده در این مقاله نشان می‌دهند که روش ترکیب عدم قطعیت مدل و پاسخ، جواب‌های بهتری را نسبت به روش‌های قبلی بر می‌گرداند، بخصوص وقتی کیفیت کار کاربران پایین باشد.

همچنین در این مقاله نشان داده شده است که فرض اینکه عدم قطعیت مدل برای داده‌های اشتباه پاسخ داده شده مجموعه آموزشی بالاست، برای مجموعه داده bmg و کلاس‌بندهای random forest استفاده شده در این مقاله فرض درستی است. در شکل ۳.۲ توزیع عدم قطعیت داده‌های درست پاسخ داده شده دارای میانگین صفر می‌باشد، در حالی که این توزیع برای داده‌های اشتباه پاسخ داده شده دارای میانگین بالایی است.



شکل ۳.۲: عدم قطعیت مدل برای داده‌های اشتباه پاسخ شده مجموعه آموزشی بالاست [۵۶].

مهم‌ترین مشکل این روش [و روش‌های قبلی] این است که چون از مجموعه داده‌های موجود و شبیه‌سازی ماشینی کاربران و خروجی‌های آنها استفاده کرده‌اند، محدودیتی روی تعداد کاربران نداشته‌اند. به عنوان مثال در بسیاری از حالت‌های که بهتر بودن یک روش از روش دیگری ادعا شده است، این ادعا بر اساس فرض وجود تعداد بسیار بالایی از کاربران انجام شده است. ممکن است این مقایسه کیفیت خروجی روش‌های مختلف وقتی که تعداد کاربران بسیار محدود باشد، دیگر معتبر نباشد، کما اینکه در نمودارها هم رسیدن به رفتار نتیجه گرفته شده پایدار معمولاً بعد از رسیدن تعداد کاربران به سطح معینی اتفاق افتاده است و قبل از آن نمودارها قابل رفتار درهمی را از خود نشان داده‌اند.

نکته مهم دیگر این است که به عنوان عدم قطعیت از احتمال نادرستی جواب حاصل از پاسخ‌های ارائه شده استفاده می‌شود، و میزان زیادی از اطلاعات بکار گرفته نمی‌شود. مثلاً می‌توان از تنوع<sup>۷</sup> یا واگرایی پاسخ‌های ارائه شده نمود (به عنوان مثال در صورتی که از مدل مخلوط گاوسی<sup>۸</sup> استفاده شود، میانگین فاصله نمونه جدید از گاوسی‌ها می‌تواند معیار خوبی برای اندازه‌گیری این تنوع باشد). یا اینکه برای هر نفر می‌توان تعداد پاسخ‌های درست و غلط وی را محاسبه و استفاده نمود ( $p_j$ های متفاوت برای کاربران مختلف) یا بین هر دو کاربری که در مورد لیبل یک مساله موافق/مخالف باشند یک لینک  $+/-$  (در حالت منفی از فرد برنده به بازنده و لینک جهت‌دار) برقرار کرد و از اطلاعات ارتباطات بین اعضا برای نظارت و بررسی یا در حالت کلی تخصیص تسک به کاربران مطلوب سیستم استفاده نمود و ....

← بررسی معیار مورد استفاده در نتایج این مقاله (مساحت زیر نمودار ROC)

<sup>۷</sup>Diversity

<sup>۸</sup>Gaussian Mixture Model (GMM)

## بهره اطلاعاتی

در [۵۵] برای مساله تصمیم‌گیری یادگیری فعال از معیار بهره اطلاعاتی برای یک مساله دو کلاسه استفاده شده است. کلاسه‌بند بکار گرفته شده در این روش یک کلاسه‌بند بر اساس مدل می‌باشد که پارامترهای آن را با  $\theta$  نمایش می‌دهیم. این پارامتر می‌تواند با روش تخمین بیشینه‌سازی احتمال پسین (MAP) از داده‌های آموزشی بدست آید. خروجی MAP یک توزیع پسین می‌باشد که  $\theta$  از آن بدست می‌آید.

یک نمونه ورودی جدید  $x$  با کلاس واقعی  $y$  در نظر بگیرید. توزیع پسین  $\theta$  را با استفاده از MAP یک بار با توجه به مجموعه آموزشی قبلی بدست آورده و یکبار با اضافه کردن داده جدید  $(x, y)$  به آن بدست آورید. این توزیع یا تغییر چندانی نمی‌کند یا تغییر آن در راستای بهبود خواهد بود (چرا که با یک داده‌ای که کلاس آن مشخص است، آن را تقویت نموده‌ایم). در مقاله فوق برای سنجش این تغییر (بهره اطلاعاتی) از اختلاف آنروپی بین دو توزیع با استفاده از روش واگرایی kullback-leibler استفاده شده است. اگر بهره اطلاعاتی بالاتر از مقدار معینی باشد، داده جدید به عنوان نمونه آموزشی مورد استفاده قرار می‌گیرد، در غیر این صورت با استفاده از کلاسه‌بند موجود دسته‌بندی می‌شود.

مشکل اصلی استراتژی فوق، این است که ما به لیبیل واقعی داده دسترسی نداریم. یکی از کارهایی که برای رفع این مشکل می‌توان انجام داد، همانند روش قبل استفاده از کلاسه‌بند فعلی برای تخمین  $y$  می‌باشد، یعنی امید ریاضی  $y$  با تابع ارزش کلاسه‌بند فعلی. دقت داشته باشیم که اگر به جای امید ریاضی از نتیجه درست/غلط کلاسه‌بندی استفاده کنیم با توجه به تعریف آنروپی و اینکه داده جدید منتج از داده‌های قبلی است، میزان بهره اطلاعاتی صفر خواهد بود. به همین ترتیب اگر احتمال تعلق به یکی از کلاس‌ها به ۱ نزدیک و دیگری به صفر نزدیک باشد، بهره اطلاعاتی بسیار پایین خواهد بود، بهره اطلاعاتی زمانی بیشینه خواهد بود که احتمال تعلق به هر دو کلاس مساوی باشد. این با انتظار ما تطابق دارد، چون وقتی کلاسه‌بند قبلی بخوبی و با احتمال بالا وضعیت یک داده را مشخص می‌کند، نیازی به بروزآوری آن نداریم و همچنین وقتی عدم قطعیت بالا می‌رود، نیاز به بروزآوری کلاسه‌بند مبرم می‌شود.

فرم ریاضی استراتژی فوق را می‌توان به این صورت بیان نمود که، اگر  $D$  مجموعه داده آموزشی فعلی بوده و  $q(\theta|D)$  تقریب فاکتوربندی شده<sup>۹</sup> احتمال پسین  $p(\theta|D)$ ، بدلیل پیچیدگی محاسبه آن باشد (استفاده از تقریب برای حفظ ادبیات بکار گرفته شده در مقاله مرجع) و با توجه به تعریف واگرایی KL  $(D_{KL}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)})$ ، می‌توان نحوه محاسبه بهره اطلاعاتی را بصورت زیر محاسبه نمود:

$$\Delta(x, y) = \mathbb{E}_{q(y|x, D)} \left[ \mathbb{E}_{q(\theta|D)} \left[ \log \frac{q(\theta|D)}{q(\theta|(x, y), D)} \right] \right] \quad (۵.۳)$$

مشکل اساسی این روش این است که اگر کلاسه‌بند با قاطعیت اشتباه کند، با این روش، ما هیچ‌گاه نسبت به اصلاح آن قدم بر نخواهیم داشت. تحلیل اینکه این اتفاق چه زمانی رخ خواهد داد به عوامل بسیاری از جمله روش اعتبارسنجی بکار گرفته شده بستگی دارد. برای جلوگیری از بروز این مشکل می‌توان از boosting چند کلاسه‌بند به جای یک کلاسه‌بند استفاده نمود (روشی نظیر روش قبلی).

## ۳.۳ اعتبارسنجی

پس از دریافت پاسخ‌های ارائه شده از طرف کاربران، نوبت به اعتبارسنجی آنها می‌رسد. در اکثر تکنیک‌های حل مساله با شبکه‌های جمعی به ازای هر سوال پیش از یک جواب دریافت می‌شود. برخی از این پاسخ‌ها ممکن است فاقد کیفیت لازم بوده و اگر در استخراج پاسخ نهایی مورد استفاده قرار گیرند، کیفیت پاسخ نهایی را پایین آورند. لذا انجام یک پیش‌پردازش برای این منظور می‌تواند موثر باشد. در این بخش ابتدا به بررسی پیش‌پردازش می‌پردازیم.

پس از انجام پیش‌پردازش، نوبت به اعتبارسنجی می‌رسد. تکنیک‌های مختلفی برای اعتبارسنجی وجود دارند. نحوه استفاده از هر کدام از این تکنیک‌ها، در مسائل مختلف، بسته به جنس مساله متفاوت است. در این بخش هم به معرفی تکنیک‌های مختلف اعتبارسنجی می‌پردازیم و هم نحوه استفاده از آنها در مسائل مختلف را بررسی می‌کنیم.

## ۱.۳.۳ پیش‌پردازش

در مرحله پیش‌پردازش، پاسخ‌هایی که دارای حداقل سطح کیفی مشخصی نباشند، حذف می‌شوند. در شبکه‌هایی نظیر MTurk می‌توان با مشخص نمودن برخی معیارها (نظیر زمانی که کاربر برای پاسخ‌گویی به سوال صرف می‌کند)، پاسخ‌های فاقد صلاحیت را مشخص نمود، تا از پرداخت هزینه به ارائه‌دهندگان آنها جلوگیری شود.

<sup>۹</sup>factorized approximation

پیش‌پردازش می‌تواند با روش‌های کاراتری، به صورت نهفته، در هنگام جمع‌آوری پاسخ‌های کاربران انجام پذیرد. به عنوان نمونه‌ای از این امر می‌توان به استفاده از سوال‌های کنترلی ساده اشاره نمود. در این روش به هنگام درخواست سوال اصلی، سوال ساده‌ای (نظیر سوالاتی که به صورت captcha در وبسایت‌های اینترنتی پرسیده می‌شود) نیز از کاربر پرسیده می‌شود، تا مطمئن شویم که کاربر به هنگام پاسخ‌گویی به سوال دقت یا هوشیاری مورد نظر را داراست. در آزمایش‌های انجام شده در [۶۴] میزان پاسخ‌های نامعتبر با استفاده از سوالات کنترلی از ۴۸.۶ درصد به ۲.۵ درصد کاهش پیدا کرده و میزان همبستگی پاسخ‌ها به پاسخ‌های افراد خیره نیز بالاتر رفته است.

از آنجایی که در سیستم‌های حل مساله جمعی، ره‌کردام از کاربران ممکن است به تعداد زیادی از مسائل از یک جنس در زمان کوتاهی پاسخ دهد؛ یک پیش‌پردازش مناسب، می‌تواند بررسی الگوهای موجود در پاسخ‌های ارائه شده توسط هر کاربر باشد. در [۶۶] نشان داده شده است که الگوی پاسخ‌های کاربران بی‌کیفیت (گزینه‌هایی که آنها در سوال‌های پشت سر هم انتخاب کرده‌اند)، دارای آنتروپی بسیار پایینی بوده است (مثلاً اکثراً یک گزینه را انتخاب کرده‌اند یا بین تعداد کمی از پاسخ‌ها مرتباً سوییچ کرده‌اند).

زمان صرف شده برای پاسخ‌گویی به مسائل نیز می‌تواند یک معیار آسان ولی مناسب دیگر باشد. در MTurk این امکان برای طراح مساله فراهم شده است که بتواند حداقل زمان مورد نیاز برای پاسخ‌گویی به مساله‌اش را تعیین نماید. در هنگام پاسخ‌گویی در صورتی که زمان صرف شده توسط کاربر از این زمان کمتر باشد، سیستم از پاسخ ارائه شده توسط وی صرف‌نظر می‌کند [۴۹].

### ۲.۳.۳ تکنیک‌های اعتبارسنجی

هر چند نحوه استفاده از تکنیک‌های اعتبارسنجی و روش‌های تحلیل آنها کاملاً وابسته به جنس مساله است، ولی قالب کلی تکنیک‌های اعتبارسنجی از چند حالت خارج نیست. مهم‌ترین این قالب‌ها عبارتند از:

**رای اکثریت.** این تکنیک مناسب مسائل کلاسه‌بندی می‌باشد. در هنگام استفاده از این تکنیک، هر کدام از مسائل به چندین عضو شبکه سپرده شده و بین جواب‌های برگشتی، یک رای‌گیری صورت می‌پذیرد. در این روش کیفیت پاسخ نهایی به دو عامل وابسته است: یکی میزان احتمال پاسخ‌گویی صحیح اعضای شبکه به سوالات، و دیگری تعداد اعضای که یک مساله به آنها سپرده می‌شود. در نتیجه مشکلات اصلی در دستیابی به یک سطح مشخص از درستی پاسخ نهایی عبارتند از: نحوه مشخص نمودن سطح کارایی افراد با توجه به سوابق آنها در سیستم، تعداد افراد مورد نیاز با توجه به کارایی‌شان جهت سپردن سوال به آنها و نحوه استنتاج پاسخ نهایی از پاسخ‌های ارائه شده توسط این کاربران. بدیهی است که برای دستیابی به یک سطح کارایی مشخص، هر چه سابقه اعضا در گذشته درخشان‌تر باشد، تعداد افراد که باید مساله به آنها سپرده شود، پایین‌تر می‌آید. فعالیت‌های متنوعی در جهت استفاده از این تکنیک در مسائل مختلف انجام شده که در بخش بعد به آنها می‌پردازیم.

**اعتبارسنجی متقابل**<sup>۱۰</sup>. یا مقایسه پاسخ‌های ارائه شده توسط کاربران مختلف با هم حالت کلی‌تری از تکنیک رای اکثریت می‌باشد. به عنوان نمونه دیگری از تکنیک‌های اعتبارسنجی متقابل می‌توان به رای‌گیری منفی/مثبت اشاره نمود که یک توسعه برای روش رای اکثریت می‌باشد و در آن کاربران امکان ارائه رای مثبت و/یا رای منفی را دارند. این توسعه در مواردی می‌تواند مفید واقع شود که یک کاربر در مواجهه با یک مساله کلاسه‌بندی چند کلاسه قادر به تشخیص کلاس درست نباشد، ولی از اشتباه بودن یک یا تعدادی از گزینه‌ها مطمئن باشد. بررسی‌های انجام شده دلالت بر این دارند، که فعالیت چندانی در زمینه رای‌گیری مثبت/منفی انجام نشده است. از دیگر روش‌های اعتبارسنجی متقابل می‌توان به استفاده از میزان همبستگی نظرات گروهی از کاربران با کیفیت ناشناس با گروهی از افراد خیره و یا کاربران دارای سوابق درخشان در سیستم اشاره نمود. در مورد همبستگی راه‌حل‌های کاربران شبکه با افراد خیره در بخش بعد بیشتر صحبت خواهیم کرد، ولی در مورد همبستگی نظرات کاربران ناشناس با کاربران با سوابق خوب، هنوز فعالیت انجام نشده است که آن را مورد بررسی قرار دهیم.

**نظرات اعضا بر فعالیت اعضا.** در این حالت یک یا چند عضو شبکه مساله را پاسخ داده و یک یا چند عضو دیگر در مورد کار پاسخ‌های ارائه شده داوری می‌کنند. این داوری می‌تواند شامل پذیرش راه حل ارائه شده، تکمیل آن و یا رد راه حل و ارجاع مساله به منبع مسائل پاسخ داده نشده باشد. مسائل مختلفی در مورد این روش می‌تواند مطرح باشد: سری یا موازی بودن داوری‌ها، تعداد داوران، نحوه انتخاب اعضای انجام دهنده مساله و داوران نمونه‌هایی از این موارد هستند. این روش برای ریز مسائلی مناسب به نظر می‌رسد که دارای پاسخ‌های تشریحی یا توصیفی باشند. فعالیت‌های چندانی در این زمینه انجام نشده است. تکنیک بررسی همبستگی نظرات کاربران ناشناس و کاربران با سوابق درخشان قرابت معنایی بالایی با این تکنیک دارد.

**تعامل و همکاری دو یا چند عضو تا رسیدن به پاسخ مساله.** در این روش کاربران با هم در ارتباط هستند تا با تعامل و همکاری با یکدیگر، در مورد پاسخ یک مساله به توافق برسند. در این تکنیک تعاملات کاربران بصورت ناشناس انجام پذیرفته و هیچ‌کدام از اعضا در مورد هویت کاربران همکار خویش

<sup>۱۰</sup> Cross validation

اطلاعاتی ندارند. پاسخ‌های ارائه شده از طرف کاربران نیز از دید بقیه همکاران آنها پنهان می‌ماند تا در پاسخ‌های ارائه شده توسط آنها تأثیر نگذارد. در این تکنیک باید روش خودکاری وجود داشته باشد که توسط آن، سیستم تشخیص دهد که کاربران به توافق نظر دست یافته‌اند یا خیر؟ تعداد اعضای تخصیص یافته به هر مساله، نحوه تعامل آنها با هم، اطلاعاتی که سیستم مجاز است بین آنها رد و بدل کند و مکانیزی که سیستم تشخیص دهد که اجماع صورت پذیرفته است، نمونه‌هایی از مسائلی هستند که در این روش باید مورد توجه قرار گیرند. فعالیت‌های انجام شده در بازی لیبیل‌گذاری تصاویر پایگاه داده گوگل تحت عنوان Google Image Labeler و پروژه پدر آن یعنی بازی ESP از مهم‌ترین فعالیت‌هایی می‌باشند که از این تکنیک استفاده نموده‌اند.

**درخواست کشف جواب‌های اشتباه از اعضای شبکه.** در این روش مسائل پاسخ داده شده به همراه کلیه پاسخ‌های ارائه شده برای هر مساله در اختیار تمامی کاربران شبکه قرار گرفته، و از آنها درخواست می‌شود تا نسبت به کشف پاسخ‌های اشتباه اقدام نمایند. مهم‌ترین مساله در این تکنیک، ایجاد انگیزه برای کاربران به منظور کشف این اشتباهات می‌باشد. فعالیت‌های چندانی در این زمینه انجام نشده است.

**تخصیص وظیفه.** در این تکنیک بر خلاف تکنیک‌های قبلی، قسمت اصلی فعالیت در ابتدا انجام می‌شود، به این صورت که مساله برای حل در شبکه قرار نمی‌گیرد، بلکه مشخصاً به فردی که صلاحیت انجام آن را دارد سپرده می‌شود<sup>۱۱</sup>. این انتخاب به قسمی صورت می‌پذیرد که احتمال درستی پاسخ ارائه شده توسط کاربر، بیشینه شود. از فعالیت‌های انجام شده در این زمینه می‌توان به [۵۸] اشاره نمود.

**بررسی سازگاری.** بررسی سازگاری فعالیت‌های فردی یک کاربر در مقایسه با هم یا حتی بررسی سازگاری پاسخ‌های ارائه شده توسط یک کاربر به یک سؤال مشخص در چندین بار تکرار پرسش در شرایط مختلف (به عنوان مثال در [۴۸]) می‌تواند به عنوان یک تکنیک اعتبارسنجی به کار گرفته شود.

**اعتبارسنجی ماشینی (خود اعتبارسنجی).** اعتبارسنجی در برخی از مسائل می‌تواند توسط خود ماشین (مثلاً با مقایسه جواب‌های دریافتی با یک جواب استاندارد کلی<sup>۱۲</sup> انجام پذیرد. به عنوان مثال، در مساله جمع‌آوری مجموعه داده‌های صوتی حاوی واژگان مشخص، برای اینکه اطمینان حاصل شود که کاربر واژه مورد نظر را ادا کرده باشد، می‌توان از چند نمونه آموزشی و یک الگوریتم هوش مصنوعی بهره برد. به عنوان نمونه‌ای از فعالیت‌های انجام شده در این زمینه می‌توان به [۴۸] اشاره نمود.

**اعتبارسنجی در لایه بالاتر.** در تکنیک‌های قبلی، اعتبارسنجی بر روی پاسخ‌های ارائه شده توسط کاربران برای یک مساله مشخص انجام می‌پذیرفت. در واقع اعتبارسنجی فقط در سطح یک مساله انجام می‌شود. در برخی از کاربردها مسائل می‌توانند طوری شکسته شوند که همپوشانی داشته باشند. در این صورت اعتبارسنجی می‌تواند در سطح پاسخ‌های بدست آمده برای مسائل مختلف همپوشان انجام پذیرد. در بررسی‌های انجام شده، فعالیت خاصی در این زمینه یافت نشد.

### ۳.۳.۳ نحوه بکارگیری

در این بخش به بررسی نحوه استفاده از اعتبارسنجی برای حل مسائل مختلف می‌پردازیم. در این قسمت مسائل کلاسه‌بندی دو کلاسه را مورد بررسی قرار می‌دهیم. مسائل کلاسه‌بندی چند کلاسه نیز با تعمیم مسائل دو کلاسه قابل حل خواهند بود. سپس برای مسائل غیر کلاسه‌بندی بیان می‌کنیم که چگونه می‌توان برخی از این مسائل را به مسائل کلاسه‌بندی تبدیل نموده و با روش‌های پیشین حل نمود. روش‌های ارائه شده برای حل مسائل خاصی نظیر مسائل نمره‌دهی و مرتب‌سازی را نیز در انتها مورد بررسی قرار خواهیم داد.

#### مسائل کلاسه‌بندی دو کلاسه

در حل مسائل کلاسه‌بندی عموماً از تکنیک رای اکثریت برای اعتبارسنجی و از اطلاعات سوابق کاربران به عنوان دانش پیشین استفاده می‌کنند. اطلاعات سوابق کاربران اغلب در قالب «احتمال ارائه پاسخ درست توسط هر کاربر» یا «کیفیت کاربر<sup>۱۳</sup>» در انجام فعالیت‌های پیشین در مساله ظاهر می‌شود. کیفیت کاربران به روش‌های مختلفی می‌تواند در مساله وارد شود. می‌توان کیفیت کار همه کاربران را یکسان در نظر گرفته (Uniform Labelers Quality) و متوسط آن را از فعالیت‌های قبلی محاسبه نمود. این احتمال را می‌توان به ازای تک تک کاربران شبکه نیز بر اساس فعالیت‌های پیشین‌شان محاسبه نمود. و یا حتی می‌توان این احتمال را به ازای هر کاربر در هر دسته از مسائل بصورت جداگانه حساب نمود. نحوه مواجهه با مساله در هر کدام از این حالت‌ها متفاوت خواهد بود. در ادامه به برخی از فعالیت‌های انجام شده در این زمینه می‌پردازیم.

در ادامه در بررسی حالت‌های مختلف از دید فنی، از نشان‌گذاری زیر استفاده می‌کنیم: مساله با  $x$  و پاسخ درست آن با  $y$  نشان داده می‌شود. پاسخ‌های کاربران

<sup>۱۱</sup> Worker filtering

<sup>۱۲</sup> Gold Standard

<sup>۱۳</sup> Labeler's Quality

مختلف را نیز با  $y_j$  نشان می‌دهیم. احتمال ارائه پاسخ درست توسط هر کاربر نیز  $p_j$  نشان داده می‌شود، یعنی  $P(y_j = y) = p_j$ . این احتمال می‌تواند برای کاربران مختلف ثابت یا وابسته به کاربر باشد. پاسخ تولید شده نهایی (برآیند پاسخ‌های کاربران دخیل در یک مساله) را نیز با  $\hat{y}$  نشان می‌دهیم. هدف کلی هر بخش نیز محاسبه کیفیت پاسخ نهایی بدست آمده<sup>۱۴</sup> یا احتمال  $q = P(\hat{y} = y)$  می‌باشد.

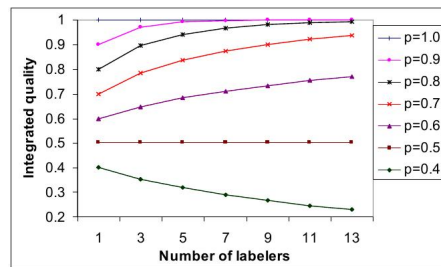
**تکنیک رای اکثریت و کاربران با کیفیت یکسان.** حل مسائل کلاسه‌بندی دو کلاسه با استفاده از تکنیک رای اکثریت و کیفیت کار یکسان برای تمامی کاربران در [۵۶] بصورت کامل مورد بررسی قرار گرفته شده است. در این مقاله همچنین کیفیت حل مساله توسط کاربران، مستقل از مسائلی که به آن‌ها واگذار می‌شود، در نظر گرفته شده است، یعنی:

$$P(y_j = y|x) = P(y_j = y) = p_j = p$$

با فرض اینکه  $2N + 1$  نفر در حل هر مساله مشارکت کنند، داریم:

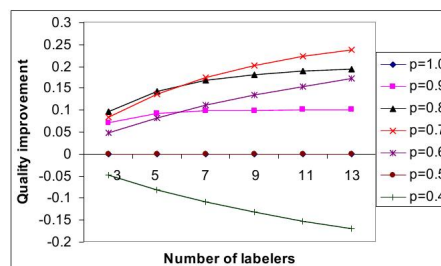
$$q = P(\hat{y} = y) = \sum_{i=0}^N \binom{2N+1}{i} p^{2N+1-i} (1-p)^i \quad (۶.۳)$$

در این رابطه،  $i$  تعداد لیبیل‌های بالقوه نادرست بوده و  $q$  احتمال این است که بیش از  $N$  نفر (حداقل  $N + 1$  نفر)، جواب درست مساله را ارائه کرده باشند. در رابطه فوق می‌توان نشان داد که تنها در صورتی که  $p > 0.5$  باشد،  $q$  بزرگ‌تر از  $p$  خواهد شد (دوگان فرض یادگیرنده‌های ضعیف در مبحث Boosting). در صورتی که  $p < 0.5$  باشد، با افزایش تعداد کاربران،  $q$  سیر نزولی طی خواهد کرد (شکل ۳.۳). همچنین اگر  $p > 0.5$  باشد، با افزایش  $N$ ، مقدار  $q$



شکل ۳.۳: ارتباط پاسخ نهایی بدست آمده برای یک مساله با تعداد کاربران مشارکت کننده در حل آن مساله، به ازای کیفیت‌های ثابت مختلف برای کاربران [۵۶].

نیز افزایش خواهد یافت، ولی نرخ تغییرات آن کاهشی خواهد بود (بهبود حاشیه<sup>۱۵</sup> با افزایش  $N$  کاهش پیدا خواهد کرد). یعنی میزان تغییرات هم به مقدار  $p$  بستگی دارد و هم به  $N$ . این مفاهیم در شکل ۳.۴ نشان داده شده‌اند. به عنوان مثال همانگونه که مشاهده می‌شود اگر  $p = 0.9$  باشد با تغییر تعداد کاربران تغییر چندانی در حاشیه صورت نخواهد پذیرفت، ولی وقتی  $p = 0.7$  باشد، افزایش تعداد کاربران، در نتیجه حاصله چشمگیر خواهد بود.



شکل ۳.۴: ارتباط نرخ بهبود در کیفیت نتیجه با تعداد کاربران مشارکت کننده در حل آن مساله، به ازای کیفیت‌های ثابت مختلف برای کاربران [۵۶].

**تکنیک رای اکثریت و کاربران با کیفیت متفاوت.** با توجه به عملکرد هر کاربر در گذشته، می‌توان احتمال ارائه پاسخ صحیح توسط وی را محاسبه نمود. در اینجا این احتمال مستقل از جنس مسائل در نظر گرفته می‌شود، یعنی  $P(y_j = y|x) = P(y_j = y) = p_j$ . می‌توان با استفاده از هیستوگرام این احتمال‌ها، نمودار دقت انجام وظایف در یک سیستم را بدست آورد.

<sup>۱۴</sup> Integrated quality

<sup>۱۵</sup> margin

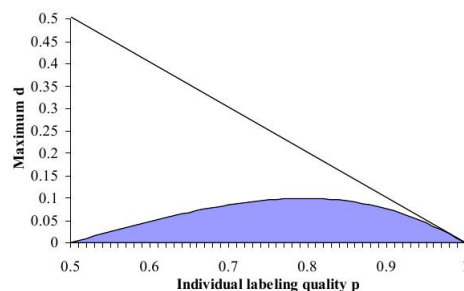
مقاله [۵۸] از کیفیت‌های متفاوت کاربران و توزیع فراوانی کیفیت همه کاربران برای حل مسائل کلاسه‌بندی استفاده نموده است. تا اول اینکه احتمال  $q = P(\hat{y} = y)$  را محاسبه نماید. محاسبه این احتمال با مساله محاسبه دقت هر کاربر در انجام مسائل قبلی، با هم یک ترکیب EM را می‌سازند. و دوم اینکه، از آنجایی که این احتمال وابسته به مقدار  $N$  (تعداد کاربران درگیر در حل مساله) است،  $N$  مناسب را برای گارانتی اینکه  $q$  از مقدار معینی بالاتر باشد، بدست آورد. یا به عبارتی ساده‌تر برای اینکه تضمین کنیم که جواب نهایی سیستم برای یک مساله مثلا دارای درستی حداقل ۹۵ درصد باشد، این مساله برای پاسخ‌گویی باید به چند کاربر سپرده شود؟

متاسفانه روابط ارائه شده در این مقاله دارای اشکالات زیادی هستند و نویسندگان مقاله به ایمیل پرسش‌های وارده بر این روابط، پاسخی نداده‌اند.

در حالت کلی نشان داده شده است [۵۶] که وقتی کیفیت همه کاربران یکسان در نظر گرفته نشود، پیچیدگی‌های زیادی در مساله بوجود خواهد آمد. حتی پاسخ‌گویی به این سوال که «استفاده از تکنیک رای اکثریت مناسب‌تر است یا رای با کیفیت‌ترین کاربر؟» مشکل خواهد بود. برای تشریح این وضعیت، مساله ساده زیر را در نظر بگیرید: در یک مساله کلاسه‌بندی دو کلاسه، از سه کاربر با کیفیت‌های  $p$ ،  $q$  و  $p + d$  بهره گرفته می‌شود. در این صورت احتمال درستی پاسخ بدست آمده با استفاده از رای اکثریت برابر خواهد بود با:

$$q = (p - d).p.(p + d) + (p - d).p.(1 - (p + d)) + (p - d).(1 - p).(p + d) + (1 - (p - d)).p.(p + d)$$

با ساده‌سازی این رابطه به مقدار  $-2p^3 + 2pd^2 + 3p^2 - d^2$  برای  $q$  می‌رسیم. از طرفی احتمال درستی پاسخ بدست آمده از طریق نظر با کیفیت‌ترین کاربر برابر خواهد بود با:  $p + d$ . اینکه کدام یک از این دو مقدار بالاترند، وابسته به مقادیر  $p$  و  $d$  می‌باشد. حل هندسی این مقایسه در شکل ۳.۵ نمایش داده شده است. در این شکل در صورتی پاسخ بدست آمده برای رای اکثریت بهتر خواهد بود که مقادیر  $p$  و  $d$  در بخش رنگی واقع شده باشند. این مثال ساده



شکل ۳.۵: مقایسه رای اکثریت و نظر با کیفیت‌ترین کاربر در یک مساله نمونه استفاده کننده از کیفیت کاربران متغیر [۵۶]. برای جزئیات بیشتر به متن مراجعه نمائید.

نشان می‌دهد که اگر کیفیت کاربران ثابت و یکسان نباشد، مساله اعتبارسنجی بسیار پیچیده خواهد شد. در صورتی که تعداد کاربران بیش از سه نفر باشد، این مساله باز هم پیچیده‌تر خواهد شد. به عنوان مثال آیا تعدادی از آرای کاربران بی‌کیفیت‌تر را حذف کنیم و سپس رای اکثریت بگیریم بهتر است یا رای اکثریت همه کاربران را یا رای کاربر با بالاترین کیفیت؟ یا .... مساله انتخاب زیر مجموعه‌ای از نظرات  $N$  کاربر، که برآیند رای اکثریت آنها بهینه باشد (دارای بالاترین احتمال درستی ممکن باشد)، می‌تواند به عنوان یک **مساله باز** جذاب مورد بررسی قرار گیرد.

**تکنیک رای اکثریت، کیفیت کاربران متغیر و دسته‌بندی مسائل.** در پژوهش‌های جدید حل مساله با جمع‌سپاری، جنس مسائلی که کاربران بر روی آنها کار می‌کنند در نظر گرفته نمی‌شود. ولی برخی پژوهش‌های پیشین در حوزه‌های مرتبط وجود دارند، که می‌توانند برای این منظور مورد استفاده قرار گیرند. به عنوان نمونه‌ای از این کارها می‌توان به [۶۷][۶۸][۶۹] اشاره نمود.

**تکنیک‌های کلاسه‌بندی نرم.** مشکل اساسی رای اکثریت این است که در آن اطلاعات میزان عدم قطعیت پاسخ ارائه شده گم می‌شود. برای این منظور می‌توان از روش‌های کلاسه‌بندی نرم<sup>۱۶</sup> استفاده نمود [۷۰]. یک نمونه از کلاسه‌بندی نرم می‌تواند اختصاص درصد پاسخ‌های درست یا هر نوع عدم قطعیت دیگر به نظر نهایی داده شده باشد.

<sup>۱۶</sup> Soft Labeling

## مسائل غیر کلاسه‌بندی

برخی از مسائل غیر کلاسه‌بندی را می‌توان از طریق تبدیل آنها به مسائل کلاسه‌بندی حل نمود. برای برخی دیگر از مسائل نیز راهکارهای مستقیمی ارائه شده است. در این بخش ابتدا نحوه تبدیل مسائل غیر کلاسه‌بندی به مسائل کلاسه‌بندی را مورد بررسی قرار می‌دهیم و سپس به حل دو مساله غیر کلاسه‌بندی نمرده‌هی و مرتب‌سازی مفهومی بصورت مستقیم خواهیم پرداخت.

**تبدیل به مسائل کلاسه‌بندی.** در [۴۶] و [۶۵] تکنیکی برای حل مسائل غیر کلاسه‌بندی از طریق تبدیل آنها به مسائل کلاسه‌بندی ارائه شده است. در این روش ابتدا یک مساله به گروهی از کاربران ارائه می‌شود و پاسخ‌های آن‌ها دریافت می‌گردد، سپس همه یا بخشی از این جواب‌ها (در صورتی که روشی برای هرس جواب‌های نامرتبطتر وجود داشته باشد) به عنوان کلاس‌های آن مساله در نظر گرفته شده و برای حل به گروه دیگری از کاربران داده می‌شوند (گروه دومی از کاربران به پاسخ‌های کاربران گروه اول رأی می‌دهند).

بدیهی است که تعداد کاربران دسته دوم باید چندین برابر کاربران دسته اول باشد، ولی تا کنون فعالیتی بر روی اینکه این دو عدد باید چه مقداری باشند تا پاسخ دارای سطح کیفی معینی باشد، انجام نشده است (**مساله باز**).

**مسائل نمرده‌هی.** در برخی از مسائل برای سنجش نظر کاربران نسبت به مساله‌ای مشخص، از آنها خواسته می‌شود تا به آن مساله، نمره‌ای در بازه‌ای مشخص ارائه دهند. در این حالت به دلیل تعدد پاسخ‌های ممکنه و همچنین عدم تفاوت بین مقادیر نزدیک به هم، استفاده از تکنیک رای اکثریت امکان‌پذیر نیست. برای مواجهه با این گونه مسائل، می‌توان از همبستگی بین نظرات کاربران<sup>۱۷</sup> و افراد خبره استفاده نمود.

در [۵۱] به منظور بررسی تکنیک همبستگی، آزمایشاتی پیشنهاد شده است. در این آزمایش‌ها همبستگی بین نظرات یک کاربر با میانگین نظرات چند کاربر گروه دیگر (مثلاً یک کاربر MTurk در مقابل چند فرد خبره و بالعکس) اندازه‌گیری می‌شود. سپس از این مقادیر (مثلاً به تعداد کاربران MTurk یا به تعداد افراد خبره) میانگین‌گیری می‌شود. مقدار نهایی بدست آمده به عنوان معیار مقایسه در نظر گرفته می‌شود (شرط معیار مقایسه این بوده است که همبستگی نهایی بدست آمده بین گروه کاربران و گروه افراد خبره، نباید از همبستگی میان گروه کاربران افراد خبره و گروه کاربران افراد خبره کمتر باشد). در صورتی که گروه افراد خبره در بین نباشد، معیار مقایسه می‌تواند عدد همبستگی بین کاربران و افراد خبره در یک تست آموزشی<sup>۱۸</sup> باشد. آزمایشات انجام شده در مقاله فوق نشان می‌دهند که یک گروه افراد خبره دارای کارایی بالاتری از یک گروه از کاربران عادی می‌باشند، در حالی که کارایی یک گروه از کاربران عادی از کارایی یک فرد خبره بالاتر می‌باشد. با استفاده از این روش همچنین می‌توان تعداد افراد مناسبی را بدست آورد، که قادرند کارایی بالاتری از یک فرد خبره بدست آورد.

در مقاله فوق از تکنیک همبستگی، برای بررسی پنج دسته فعالیت در حوزه پردازش زبان طبیعی استفاده شده است (این فعالیت‌ها عبارتند از: تحلیل عواطف موجود در متن، شباهت جملات از لحاظ معنایی، بازشناسی استلزام متنی، تشخیص رخداد‌های زمانی و تشخیص معنای کلمات مبهم). در هر فعالیت چندین مساله طراحی شده‌اند. این مسائل ابتدا برای حل به افراد خبره داده شده و سپس توسط جمعی از کاربران شبکه MTurk حل شده‌اند. نتایج بدست آمده قابل توجه می‌باشند. به عنوان مثال، در فعالیت تحلیل عواطف، کاربران در مورد پنج حس از شش حس (به جز حس ترس) توانسته‌اند به کارایی بالاتر از نظر یک فرد خبره دست یابند. حداقل تعداد افراد لازم برای اینکه درستی فعالیت این کاربران از یک فرد خبره کمتر نباشد، بین ۲ تا ۹ نفر برای حس‌های مختلف متفاوت بوده است (خشم-۲، انزجار-۲، لذت-۷، ناراحتی-۲ و تعجب-۹). در بررسی شباهت معنایی جملات نیز، همبستگی نظرات ۱۰ نفر کاربر تقریباً نزدیک نظر یک فرد خبره بوده است. همچنین در تمامی فعالیت‌ها، نتیجه نهایی کار با استفاده از تعداد کاربران بیشتر بالاتر می‌رفته است (نمودار همبستگی نظرات یا درستی فعالیت بر حسب تعداد کاربران مشارکت کننده صعودی بوده است).

**مسائل مرتب‌سازی (یا ارزش‌گذاری چندین گزینه).** این مساله در مقاله [۴۷] به منظور مرتب‌سازی مفهومی تعدادی تصویر بر اساس یک ویژگی مفهومی مشخص مطرح شده است. برای این منظور الگوریتم CollaboRank معرفی شده است که در آن به هر کدام از کاربران، تعدادی جفت تصویر ارائه می‌شود تا نسبت به مقایسه آن‌ها با توجه به معیار مورد نظر مساله اقدام نمایند. نتیجه مقایسه در گرافی که هر کدام از راس‌های آن یکی از تصاویر می‌باشند ذخیره می‌شود، به این صورت که اگر تصویر الف با تصویر ب مقایسه شود، یالی از الف به ب کشیده می‌شود. وزن این یال هر چه از ۰.۵ به سمت ۱ بیشتر باشد، نشان دهنده ارجعیت تصویر مبدأ یال می‌باشد و هر چه از ۰.۵ به سمت صفر کمتر باشد، نشان دهنده ارجعیت تصویر مقصد یال است. وزن ۰.۵ نیز نشان دهنده عدم ارجعیت است. در انتهای مقایسه‌ها، برای هر گره درجه خروجی آن منهای درجه ورودی آن محاسبه شده و گره با بیشترین وزن به عنوان گره با بالاترین ترجیح انتخاب شده و همراه با تمامی یال‌هایش از گراف خارج می‌شود. همین عملیات برای استخراج گره‌های با ترجیح بدی نیز تکرار می‌شود تا زنجیره مرتب شده تصاویر حاصل شود.

تنها نکته باقیمانده در مورد این تکنیک، نحوه انتخاب تصاویری است که برای مقایسه در اختیار کاربر قرار می‌گیرد. در هر مقایسه  $x$  تصویر در اختیار کاربر

<sup>۱۷</sup>interannotator agreement

<sup>۱۸</sup>Gold Standard



قرار می‌گیرند. به این صورت که ابتدا تصویر با بالاترین آنتروپی انتخاب می‌شود (گره با بالاترین مجموع وزن بال‌های اتصالی) و سپس  $x - 1$  تصویر دیگر به قسمی انتخاب می‌شوند که دارای بیشترین فاصله‌ها از تصویر انتخاب شده قبلی باشند (دارای یالی با بیشترین وزن باشند).

در یک آزمایش با ۲۱ نفر کاربر و ۵ کلاس ۵۰ تصویری، کاربران در طی مجموعاً ۲۰ دقیقه (متوسط زمان هر کاربر برای هر کلاس از تصاویر بین ۱۱ تا ۱۷ ثانیه) با انجام ۸۶۸۸ مقایسه اقدام به مرتب‌سازی تصاویر هر کلاس نمودند. ضریب همبستگی جواب‌های ارائه شده از کاربران مختلف در کلاس‌های متفاوت این آزمایش‌ها مقادیر بزرگی بوده (از ۰.۵۱ تا ۰.۷۵) و نشان از هم‌رای بودن کاربران در مرتب‌سازی‌ها می‌دهد. نتایج مرتب‌سازی نیز از لحاظ مفهومی دارای کیفیت بالایی می‌باشند، که با توجه به نتیجه حاصله قبلی، دور از انتظار نبوده است.

**سایر مسائل.** مسائل مختلف دیگری نیز در این زمینه وجود دارند که کمتر به آنها پرداخته شده است. یکی از این موارد نحوه مدیریت مسائلی است که دارای وابستگی‌های پیش‌نیازی هستند، که فعالیت خاصی در مورد آنها انجام نشده است.

یکی دیگر از این مسائل، نحوه اجرای ریز وظایفی است که در سیستم نیاز به هم‌زمانی دارند (مثلاً تعداد مشخصی از کاربران باید بصورت هم‌زمان به انجام مجموعه‌ای از وظایف بپردازند). به عنوان نمونه‌ای از کارهای انجام شده در این زمینه می‌توان به [۶۳] اشاره نمود. البته راهکار ارائه شده در این مقاله برای مسائلی مناسب است که تعداد کاربران هم‌زمان مورد نیاز آن‌ها از تعداد محدودی بالاتر نباشد. تحلیل اینکه این تعداد چقدر است و چگونه می‌توان رویکردی برای تعداد افراد هم‌زمان مورد نیاز بیشتری ارائه کرد، هنوز باز است.

مبحث دیگری که به آن پرداخته نشده است، نحوه مدیریت کارای مسائلی است که برخط بوده و مدت زیادی نمی‌توانند در صف انتظار برای پاسخگویی بمانند. یک مبحث باز دیگر در این زمینه، مدل‌سازی عملکرد کاربران می‌باشد. همانگونه که دیدیم با استفاده از تکنیک رای اکثریت، یک مساله برای حل، حداقل باید به سه نفر سپرده شود، برای تضمین سطح مشخصی از درستی پاسخ، این عدد بسیار بالاتر می‌رود. با تکنیک‌های مدل‌سازی و یادگیری عملکرد کاربران می‌توان انتظار داشت که این سربار به مقدار زیادی کاهش پیدا کند.

## ۳.۴ سایر مباحث مرتبط

در این بخش به بررسی یک مساله دیگر می‌پردازیم که در دسته‌بندی فوق قرار نمی‌گیرد. این مبحث تصحیح بایاس کاربران می‌باشد. در بسیاری از مسائل، بخصوص مسائل نظرسنجی، یک اصل خیلی مهم این است که کاربرانی که به سوالات پاسخ می‌دهند، نمونه خوبی از جامعه هدف مورد نظر باشند. ولی در شبکه‌های جمعی، بنا به دلایل بدیهی، عمدتاً این اصل نمی‌تواند برقرار باشد. سوالی که اینجا می‌تواند مطرح باشد این است که آیا می‌توان نسبت به تصحیح بایاس کاربران در این شبکه‌ها اقدامی را انجام داد یا نه؟ این مساله در این بخش مورد بررسی قرار خواهد گرفت.

## ۱.۳.۴ تصحیح بایاس کاربران

یکی از ایراداتی که به کاربرد شبکه‌های جمعی در حل برخی از مسائل (نظیر ارزیابی‌ها و نظرسنجی‌ها) وارد می‌شود، بایاس بودن آن‌هاست (بدلیل اینکه جمعیت موجود در شبکه نمونه عادلانه‌ای از کل جمعیت مورد مطالعه نیست). هر چند در مورد مدل‌های نمونه‌گیری قدیمی، روش‌هایی برای تصحیح این بایاس‌ها در نمونه‌های نابایاس از جمعیت‌های هدف ارائه شده‌اند (نظیر [۶۱][۶۲][۵۲][۵۳][۵۴])، ولی هنوز فعالیت چندانی در این زمینه برای شبکه‌های جمع‌سپاری جدید انجام نگرفته است.

از معدود کارهای انجام شده در این زمینه می‌توان به [۵۱] اشاره نمود، که در آن یک رهیافت مبتنی بر Gold-standard و استفاده از یک مجموعه آموزشی لیبل‌دار برای مسائل کلاسه‌بندی مورد بررسی قرار گرفته است. که در ادامه به آن می‌پردازیم.

یک مساله دو کلاسه با مسائل  $x_i$  و کلاس‌های واقعی  $y_i$  در نظر بگیرید. برای هر مساله  $x_i$ ، نظر  $w$  کاربر گرفته می‌شود  $(y_{i1}, \dots, y_{iw})$ . برای پیش‌بینی مساله  $x_i$ ، با استفاده از MAP می‌توان بصورت زیر عمل کرد:

$$\text{Likelihood} : L = \log \frac{P(y_i = + | y_{i1}, \dots, y_{iw})}{P(y_i = - | y_{i1}, \dots, y_{iw})} \implies C = \begin{cases} + & \text{if } L > 1 \\ - & \text{if } L < 1 \end{cases} \quad (۷.۳)$$

$$L = \log \frac{P(y_{i1}, \dots, y_{iw} | y_i = +) P(y_i = +)}{P(y_{i1}, \dots, y_{iw} | y_i = -) P(y_i = -)}$$

از طرفی می‌دانیم  $P(y_{i1}, \dots, y_{iw} | y_i = +) = \prod_{j=1}^w p(y_{ij} | y_i = +)$  پس:

$$\log L = \sum_{m=1}^w \log \frac{P(y_{im}|y_i=+)}{P(y_{im}|y_i=-)} + \log \frac{P(y_i=+)}{P(y_i=-)}$$

برای این منظور می‌بایست احتمال‌های  $P(y_i = L)$  و  $P(y_{ij}|y_i = L)$  را بدست آورد. این مقادیر را می‌توان با استفاده از یک مجموعه داده آموزشی دارای لیبل بدست آورد.  $P(y_i = L)$  در ساده‌ترین حالت برابر احتمال‌های پیشین خواهند شد و  $P(y_{ij}|y_i = L)$  را می‌توان با استفاده از تخمین ML بدون Laplace smoothing از ماتریس Confusion سوابق یک فرد استخراج نمود.

تخمین MAP به جای ML در معادلات فوق، معادل قانون رای‌گیری وزن‌دار خواهد شد، که وزن رای هر کاربر متناسب با نسبت log likelihood پاسخ‌های ارائه شده وی در مجموعه آموزشی خواهد بود.

## منابع و مراجع

- [1] Jeff Howe, "The Rise of the Crowdsourcing", Wired Magazine, 2006 , [link](#).
- [2] Jeff Howe, "Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business", Crown Business publication, 2008.
- [3] "The DARPA Network Challenge", Project Report, 2010.
- [4] T. W. Malone, "What is collective intelligence and what will we do about it?", Edited transcript of remarks at the official launch of the MIT Center for Collective Intelligence, October 13, 2006.
- [5] D. Deutsch, "Quantum Theory, The Church-Turing principle and the universal quantum computer", Proc. R. Soc. Lond. A 400, 97-117, 1985.
- [6] H. G. Rice, "Classes of recursively enumerable sets and their decision problems", Trans. AMS 89:25-59, 1953.
- [7] D. Hofstadter, "Göedel, Escher, Bach", Basic Books, New York, 1979.
- [8] M. L. Minsky, "Future of AI Technology", Toshiba Review, Vol.47, No.7, July 1992.
- [9] G. Chaitin, "Meta Math!: The Quest for Omega", Pantheon Books, New York, 2005.
- [10] D. Prelec, "A Bayesian Truth Serum for Subjective Data", Science, pp. 462-466, 2004.
- [11] A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungs problem", Proc. London Math. Soc. Ser. 2 42, 230-265, 1936; A correction, ibid, 43, 544-546, 1937.
- [12] Y. Matiyasevich, J. P. Jones, "Direct translation of register machines into exponential Diophantine equations", In L. Priese, editor, Report First GFI Workshop Found. Theor. Computer Sci., pages 117-130, Univ. Gesamthochschule, Paderborn, 1983.
- [13] E. Zermelo, "Über Grenzzahlen und Mengenbereiche", Fund. Math. 16, 29-47, 1930.
- [14] E. von Hippel, "Sticky information and the locus of problem solving: Implications for Innovation", Management Science 40(April), 429-439, 1994.
- [15] R. Penrose, "The Emperor's New Mind", Oxford Press, Oxford, U.K., 1989.
- [16] J. McCarthy, "Review of The Emperor's New Mind by Roger Penrose", Bulletin of the American Mathematical Society, v23 i2, 606-616, 1998.
- [17] R. Penrose, "The Shadows of the Mind", Oxford: Oxford University Press, 1994.

- [18] "Handbook of Collective intelligence", A collective Intelligence Wiki book, [link](#).
- [19] A. Church, "A Note on the Entscheidungsproblem", *Journal of Symbolic Logic* (1): 40–41, 1936.
- [20] A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem", *Proceedings of the London Mathematical Society* 42: pp. 230–265, 1937.
- [21] N. Dershowitz, Y. Gurevich, "A natural axiomatization of computability and proof of Church's Thesis". *The Bulletin of Symbolic Logic*, 14(3):299-350, 2008.
- [22] U. Boker, N. Dershowitz, "The Church-Turing Thesis over Arbitrary Domains", *Pillars of Computer Science: Essays Dedicated to Boris (Boaz) Trakhtenbrot on the Occasion of His 85th Birthday*, Arnon Avron, Nachum Dershowitz, and Alexander Rabinovich, eds., *Lecture Notes in Computer Science*, vol. 4800, Springer-Verlag, Berlin, pp. 199–229, 2008.
- [23] U. Boker, N. Dershowitz, "Three Paths to E $\square$ ectiveness", *Fields of Logic and Computation: Essays Dedicated to Yuri Gurevich on the Occasion of His 70th Birthday*, Andreas Blass, Nachum Dershowitz, and Wolfgang Reisig, eds., *Lecture Notes in Computer Science*, vol. 6300, Springer-Verlag, Berlin, 2010.
- [24] Marcin Milkowski, "Is computationalism trivial?", Gordana Dodig Crnkovic and Susan Stuart (eds.), *Computation, Information, Cognition – The Nexus and the Liminal*, Cambridge Scholars Publishing, 2007, pp. 236-246.
- [25] Bickhard, M. H. (forthcoming). Troubles with Computationalism. In R. Kitchener, W. O'Donohue (Eds.) *Psychology and Philosophy: Interdisciplinary Problems and Responses*, ????
- [26] Richard Samuels, "Classical Computationalism and the Many Problems of Cognitive Relevance", *Studies in History and Philosophy of Science*, ????
- [27] Gualtiero Piccinini, *Computationalism in the Philosophy of Mind*, *Philosophy Compass* 4 (2009).
- [28] Gualtiero Piccinini, "Computationalism, The Church-Turing thesis, and the Church-Turing fallacy",???
- [29] Chalmers, D. J. (1995), "Facing up to the problem of consciousness," *Journal of Consciousness Studies*, vol. 2, pp. 200–219.
- [30] Seyedreza Maleeh, "The Conscious Mind Revisited: An Informational Approach to the Hard Problem of Consciousness", Doctor of Philosophy thesis, University of Osnabrück, ???
- [31] Drew McDermott, "Artificial Intelligence and Consciousness", ???
- [32] Dannett, The practical requirements for making a conscious robot,???
- [33] Tovey, M., editor, *Collective Intelligence: Creating a Prosperous World at Peace*. Oakton, VA: EIN Press. 648 pp.
- [34] Giorgio Buttazzo, *Artificial Consciousness: Utopia or Real Possibility?*, 2001
- [35] R. Kurzweil, *The Age of Spiritual Machines*, Viking Press, New York, 1999.
- [36] G.S. Paul and E. Cox, *Beyond Humanity: CyberEvolution and Future Minds*, Charles River Media, Rockland, Mass., 1996.

- [37] H. Moravec, *Robot: Mere Machine to Transcendent Mind*, Oxford University Press, Oxford, UK, 1999.
- [38] Olivier Bourneza, Nachum Dershowitz, "Foundations of Analog Algorithms",???
- [39] Gualtiero Piccinini, *Computationalism in the Philosophy of Mind*, *Philosophy Compass* 4 (2009).
- [40] Matt Careter, *Mind and Computers: An Introduction to the philosophy of artificial intelligence*, ???
- [41] John McCarthy, *From here to human-level AI*,????
- [42] Dreyfus, Dreyfus, *Mind Over Machine*, ???
- [43] David L. Waltz, "Artificial Intelligence", *Scientific American*, October 1982.
- [44] Howe, J. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. Crown Business, New York, 2008.
- [45] T. W. Malone, *The Future of Work*. Cambridge, MA: Harvard Business School Press, 2004.
- [46] M. S. Bernstein, G. Little, R. C. Miller, B. Hartmann, M. Ackerman, D. Karger, D. Crowell, K. Panovich, "Soylent: A Word Processor with a Crowd Inside", *Proceedings of UIST 2010: ACM Symposium on User Interface Software and Technology*. New York, NY, 2010.
- [47] J. H. M. Janssens, "Ranking Images on Semantic Attributes using Human Computation", In *NIPS Workshop on Computational Social Science and the Wisdom of Crowds*, Whistler, Canada, December 2010.
- [48] J. Ledlie, B. Odero, E. Minkov, I. Kiss, J. Polifroni, "Crowd Translator: On Building Localized Speech Recognizers through Micropayments", *SIGOPS Operating Systems Review*, Volume 43, Issue 4, 2009.
- [49] W. Mason, S. Suri, "Conducting Behavioral Research on Amazon's Mechanical Turk", *Social Science Research Network Working Paper Serie*, Oct. 2010.
- [50] J. Yang, L. A. Adamic, and M.S. Ackerman, "Crowdsourcing and knowledge sharing: strategic user behavior on taskcn", in *Proc. ACM Conference on Electronic Commerce*, pp.246-255, 2008.
- [51] R. Snow, B. O'Connor, D. Jurafsky, and A.Y. Ng, "Cheap and Fast - But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks", in *Proc. EMNLP*, pp.254-263, 2008.
- [52] A. P. Dawid and A. M. Skene. "Maximum Likelihood Estimation of Observer Error-Rates Using the EM Algorithm", *Applied Statistics*, Vol. 28, No. 1, pp. 20-28, 1979.
- [53] J. M. Wiebe, R. F. Bruce, T. P. O'Hara, "Development and use of a gold-standard data set for subjectivity classifications", In *Proc. of ACL-1999*.
- [54] P. S. Albert, L. E. Dodd, "Cautionary Note on the Robustness of Latent Class Models for Estimating Diagnostic Error without a Gold Standard", *Biometrics*, Vol. 60, pp. 427-435, 2004.
- [55] U. Paquet, J. V. Gael, D. Stern, G. Kasneci, R. Herbrich, T. Graepel, "Vuvuzelas Active Learning for Online Classification", *Computational Social Science and the Wisdom of Crowds Workshop (colocated with NIPS10)*, 2010.
- [56] V.S. Sheng, F.J. Provost, and P.G. Ipeirotis, "Get another label? improving data quality and data mining using multiple, noisy labelers", in *Proc. KDD*, pp.614-622, 2008.

- [57] P. Ipeirotis, F. Provost, V. Sheng, J. Wang, "Repeated Labeling Using Multiple, Noisy Labelers", Working paper series, CeDER-10-03, 2010.
- [58] N. Eagle, "txteagle: Mobile Crowdsourcing", Lecture Notes in Computer Science: Internationalization, Design and Global Development, vol 5623, Springer Berlin / Heidelberg, pp 447-456, 2009.
- [59] L. B. Chilton, J. J. Horton, R. C. Miller, S. Azenkot, S., "Task search in a human computation market", In Proceedings of the ACM SIGKDD workshop on human computation, pp. 1-9, 2010.
- [60] C. F. Camerer, R. M. Hogarth, "The effects of financial incentives in experiments: A review and capital-labor-production framework", Journal of Risk and Uncertainty, 19, 7-42, 1999.
- [61] R. A. Berk, "An introduction to sample selection bias in sociological data", American Sociological Review, 48(3), 386-398, 1983.
- [62] J. J. Heckman, "Sample selection bias as a specification error", Econometrica, 47(1), 153-161, ficial cultural market. Science, 311(5762), 854-856, 1979.
- [63] S. Suri, D. J. Watts, "Cooperation and contagion in networked public goods experiments", 2010. link: <http://arxiv.org/abs/1008.1276>
- [64] A. Kittur, E. H. Chi, B. Suh, "Crowdsourcing user studies with Mechanical Turk", In Proceedings of the ACM conference on human factors in computing systems, pp. 453-456, 2008.
- [65] G. Little, L. B. Chilton, M. Goldman, R. C. Miller, "Exploring iterative and parallel human computation processes", In Proceedings of the ACM SIGKDD workshop on human computation, pp. 68-76, 2010.
- [66] D. Zhu, B. Carterette, "An analysis of assessor behavior in crowdsourced preference judgments", In Proceedings of the ACM SIGIR workshop on crowdsourcing for search evaluation, pp. 1-6, 2010.
- [67] A. P. Dawid, A. M. Skene, "Maximum likelihood estimation of observer error-rates using the EM algorithm", Applied Statistics 28-1, 20-28, 1979.
- [68] P. Smyth, "Bounds on the mean classification error rate of multiple experts", Pattern Recognition Letters 17-12, 1996.
- [69] P. Smyth, U. M. Fayyad, M. C. Burl, P. Perona, P. Baldi, "Inferring ground truth from subjective labelling of Venus images", In NIPS94, pp. 1085-1092, 1994.
- [70] P. Smyth, "Learning with probabilistic supervision", In Computational Learning Theory and Natural Learning Systems, Vol. III: Selecting Good Models, T. Petsche, Ed. MIT Press, 1995.
- [71] A. Gelman, J. B. Carlin, H. Stern, D. B. Rubin, "Bayesian Data Analysis", 2nd ed. Chapman and Hall/CRC, 2003.
- [72] ارنست ناگل، ج نیومن و آلفرد تارسکی، «برهان گودل و حقیقت و برهان»، ترجمه محمد اردشیر، تهران، انتشارات مولی، چاپ اول، ۱۳۶۴.
- [73] جعفر محمدی، «شبکه‌های اجتماعی و قدرت اطلاعاتی آنها»، گزارش تکنیکی، آزمایشگاه رسانه دیجیتال دانشگاه صنعتی شریف، تابستان ۱۳۸۹.