1. من اول یه تشکر بکنم از دکتر حالا چون ارائه اخر هست و جلسه اخر و به نظرم درس خیلی اینتراکتیو بود. از بقیه درس ها که فقط استاد لکچر میداد بیشتر یاد گرفتیم.
2. اول یه چیز جالب در مورد این مقاله بگم که طبق گفته خودشون ۳ جا هست. هم IEEE . هم CVPR هم ICML. اول IEEE گذشته بعد ورک شاپ.
3. هیچی
4. خب مقاله ما در مورد رانندگی خودکار شهری هست. رانندگی خودکار الان مدت هاست توسط شرکت های بزرگ گوگل تسلا و غیره سرمایه گذاری میشه. کلی کار انجام شده و هنوز هم خیلی جای کار داره. این مقاله یه قسمتی از این دریای بزرگ رو میخواد انجام بده. یه نکته ای هم بگم که اینها هیچ کدوم بدون باگ نیستن. مقاله رو خوب پرزنت کردن و ارائه دادن کار ارزشمندی هم هست ولی اگر سر به گیت‌هابش بزنید کلی خطا داره و مشکلات حل نشده هست. البته فیلمش هم موجوده مهدی بهمون نشون میده.
5. اگر بخوام مقاله رو در یک اسلاید توضیح بدم که کلیتش دستتون بیاد. ما در اینجا دوست داریم با روش‌های یادگیری تقویتی عمیق + مدل latent یک ماشین خودکار درست کنیم. یک تفاوت اساسی این مقاله کلمه end-to-end است یعنی فقط یه ماژول خاص رو نداره بلکه کل ماژول های perpception decision یا تصمیم و کنترل رو در برمیگره. به علاوه این که این مدل تفسیرپذیر است و حتی بتونه در مدل های پیچیده تر شهری هم کار کنه. در نهایت هم میخوایم بگیم که مدل از یه سری مدل های بیس لاین دیگه بهتر عمل میکنه.
6. اگر هم بخوام contribution اصلی مقاله رو بگم. میتونم بگم یک فضایی به نام Latent space معرفی کرده که در واقع چیزهایی که در مورد محیط یاد میگیره رو اونجا encode میکنه. حالا شامل فیچرهای مکانی، اطلاعات جاده، اطلاعات کاربران. در واقع اطلاعات قدیمی که از محیط high-resolution یاد گرفته اینجا بصورت فشرده در یه فضای با ابعاد خیلی کمتر موجوده که حالا حسین خیلی بهتر برامون توضیح میده که قضیه از چه قراره.
7. همون طور که گفتم اکثر روش های موجود از مدل های به شدت ماژولارایز شده استفاده میکنند یعنی کاملا جدا هستند از هم. مثلا فرایند دید رو در نظر بگیرید. تعداد بسیار بسیار زیادی فیچرهایی که بصورت دستی extract شدند هست و خب این در فضاهای limited کار میکنه ولی در فضاهای شهری پیچیده تر نه. علتش هم دو چیز هست. یکی این که این روش ها از هیوریستیک های بیش از حد استفاده میکنند که باعث میشه رانندگی خیلی محتاطانه بشه و دوم این که خیلی سخته که این شهود ها رو برای تک تک تسک های مختلف رانندگی تعریف کنیم. راه حل؟ راهکار اینه که از این روش ها قدری دوری کنیم و روش هایی رو اپلای کنیم که سیاست یا Policy یاد گرفته بشه. علاوه بر این همین باعث میشه که بتونیم به توانایی خارق العاده superhuman برسیم که خب با روش های معمولی نمیشد.
8. کلا دو تا روش داریم: یکی روش های بر مبنای تقیده. Imitation. یعنی تقلید از یه سری راننده حرفه ای حین رانندگی. روش های RL هم که مشخصه. میخوایم بر اساس کاوش و یه نوعی سیستم پاداش خطا سیاست ها رو یادبگیریم. ولی ولی دو مشکل اساسی داریم که مشخصا روش هامون تفسیرپذیر نیستند و دومین مشکل هم اینه که خب فقط یه سری کارها مثل حرکت بین لاین خیابون و یا مثلا دنبال یه ماشین رفتن رو یاد میگیره که خب خیلی ساده اس. ترافیک اصولا یه سیستم پویا هست و خیلی تعامل زیادی نیاز داره. شرایط مختلف و غیره خیلی کار رو سخت میکنه که بشه خوب یاد گرفت.
9. روش یادگیری تقلیدی رو گفتم. پس یه سری متخصص به عنوان داده لیبل دار تسک یادگیری رو راحت میکنه. معمولا راحت هم ترین میشن این روش ها و خوب کار میکنه اگر بتونیم داده خوب جمع کنیم و خب همین مشکل اصلیش هست که دیتا هانگری هست و دیتا انقدری که نیاز داره موجود نیست. یه مشکل مهم دیگه هم که داره اصلا یه سری جاها رو نمیشه دیتا جمع کرد. مثلا داده وقتی که تو با کامیون رو به رو میشیم یا مثلا تو دره در حال سقوطیم رو نداریم و این روش های هیچ ایده ای ندارن مطلقا.
10. روش های RL هم که اخیرا خیلی خوب عمل کردن و مشخصا deep RL تونسته حتی فراتر از انسان عمل کنه در مثلا بازی go که گوگل چندسال با پروژه alphago تونست بهترین بازیکن انسانی رو شرکت بده. این هم ساختمان deep mind هست که تو بازی اتاری با همین روش از تونست خیلی خوب عمل کنه. خوبی RL اینه که میتونه کیس های ممکن رو (حتی خطرناک ها رو ) بدون دیتای متخصص یاد بگیره. مشکل اصلی شون هم همین طور که گفتم تفسیرپذیریه که مشخصا میدونیم روش های دیپ این مشکل رو دارند. خیلی مهمه که بدونیم چگونه ماشین محیط رو میفهمه و ایا اصلا درک درستی داره یا نه. تو این مقاله روشمون محیط رو به درستی درک میکنه و تفسیرپذیر هست.