

طرح درس جهت ارائه در نیمسال تحصیل دوم

دانشکده	مهندسی برق و کامپیوتر	گروه	مهندسی پزشکی
گرایش	مهندسی پزشکی	مقطع	کارشناسی ارشد/دکتری
نام درس	مباحث ویژه-استنتاج علی Causal Inference 01_6606092	نوع درس	پایه ☐ نظری ☒ تخصصی ☒ عملی ☐ اختیاری ☒ نظری-عملی ☒
تعداد واحد	۳	نام استاد	منصوره پاکروان
دروس پیش‌نیاز	یادگیری ماشین یا تشخیص آماری الگو یا شبکه های عصبی	تلفن دفترکار	۸۲۸۸۴۹۷۲
دروس هم‌نیاز		پست الکترونیک	mpakravan@modares.ac.ir

✓ اهداف درس:

- ✓ آشنایی با مفاهیم نردبان علیت
- ✓ توانایی کار با گراف های علی و اندازه گیری اثر علی درمان روی خروجی
- ✓ توانایی کار با داده ها به منظور استخراج گراف علی
- ✓ توانایی کار با زبان برنامه نویسی پایتون و کتابخانه های موردنظر برای استنتاج علی
- ✓ آشنایی با مفاهیم تفسیر پذیری و توضیح پذیری در یادگیری ماشین
- ✓ توانایی کار با زبان برنامه نویسی پایتون و کتابخانه های موردنظر برای تفسیر مدل های جعبه سیاه

✓ رئیس مطالب و برنامه ارائه در کلاس: (در صورتی که واحد عملی یا نظری-عملی بود، نوع آموزش در توضیحات بیان شود)

شماره جلسه	موضوع جلسه درس	توضیحات
جلسه اول	• چرا استنتاج علی؟	
جلسه دوم	• مفاهیم پایه ○ مداخله ○ گیج کننده (Confounding) ○ پارادوکس سیمپسون ○ همبستگی های کاملاً جعلی • نردبان علیت	
جلسه سوم	• پله اول نردبان علیت: آمار و احتمال ○ قوانین احتمال ○ مقدار مورد انتظار، واریانس و کوواریانس ○ رگرسیون خطی	
جلسه چهارم	• مدل های ساختاری علی (Structural Causal Model) ○ مفاهیم پایه ای گراف	
جلسه پنجم	• مدل های گرافی ○ استقلال در توزیع و استقلال در گراف ○ فرض مارکوف محلی	

جلسه ششم	- واحدهای سازنده مدل های گرافی ○ زنجیره، دوشاخه، برخورددهنده ○ استقلال شرطی
جلسه هفتم	- گراف بدون دور جهت دار (DAG) ○ متغیرهای مورد نیاز DAG
جلسه هشتم	- قوانین d-جدایی (d-separation) - تستهای استقلال شرطی
جلسه نهم	- پله دوم نردبان علیت: مداخله (Intervention) - جریان ارتباط علی و جریان علیت
جلسه دهم	- آزمایش تصادفی کنترل شده - مداخله برای تعیین اثر علی (Causal effect) ○ شرط کردن ○ فرض مدولاریتی ○ تعریف اثر علی
جلسه یازدهم	پله سوم نردبان علیت: خلاف واقعیت (Counterfactual)
جلسه دوازدهم	- متوسط اثر درمان (Average Treatment Effect) - فاکتورسازی کوتاه شده (truncated) - تعدیل کردن (Adjustment)
جلسه سیزدهم	- معیار درب پشتی (Backdoor criteria) - تعدیل اثر علی
جلسه چهاردهم	معیار درب جلویی (Frontdoor criteria)
جلسه پانزدهم	محاسبات عملگر do
جلسه شانزدهم	تفکیک اثر علی مستقیم از غیرمستقیم
جلسه هفدهم	- استنتاج علی در سیستم های خطی ○ رگرسیون خطی در مقابل معادلات SCM - متغیر ابزاری (Instrumental variable)
جلسه هجدهم	- کلاس هم ارزی مارکف - ارزیابی مدل علی گرافی ○ آزمون های ابطال
جلسه نوزدهم	- مدل های علی، فرض ها و چالش ها ○ قابلیت شناسایی ○ فرض مثبت بودن احتمال ○ فرض ماژولار بودن ○ فرض سازگاری ○ فرض برون زایی متغیرهای کمکی ○ فرض ارزش واحد درمان پایدار ○ فرض مبادله پذیری
جلسه بیستم	- الگوریتم های کشف گراف علی ○ الگوریتم PC ○ الگوریتم LiNGAM
جلسه بیست و یکم	متوسط اثر علی شرطی (CATE)
جلسه بیست و دوم	وزن دهی معکوس احتمال (Inverse Probability Weighting)
جلسه بیست و سوم	یادگیری ماشین و استنتاج علی

	- متوسط اثر علی شرطی - معیار های ارزیابی نتایج استنتاج علی	
جلسه بیست و چهار	- فرا یادگیرنده ها (Meta learners) - فرا یادگیرنده S-Learner - فرا یادگیرنده T-Learner - فرا یادگیرنده X-Learner - فرا یادگیرنده دوجندان قوی (Doubly Robust)	
جلسه بیست و پنجم	- فرا یادگیرنده ها (Meta learners) - فرا یادگیرنده یادگیری ماشین دوگانه (DML) - TARNet - DragonNet - فرا یادگیرنده SNet (Shared Net)	
جلسه بیست و ششم	- یادگیری ماشین تفسیرپذیر: مقدمه - مدل های جعبه سفید و تفسیر آن - رگرسیون خطی - درخت تصمیم - جنگل تصادفی - مدل های افزودنی تعمیم یافته (Generalized additive models)	
جلسه بیست و هفتم	- مدل های جعبه سفید و تفسیر آن - الگوریتم تقویت گرادیان (Gradient Boosting) - ماشین تقویت قابل توضیح (Explainable Boosting Machine)	
جلسه بیست و هشتم	- روش های تفسیر محلی مستقل از مدل - روش LIME (Local Interpretable Model-agnostic Explanations)	
جلسه بیست و نهم	- روش های تفسیر محلی مستقل از مدل - روش SHAP (SHapley Additive exPlanations)	
جلسه سی ام	- روش های تفسیر مدل برای شبکه های عصبی کانولوشن - روشهای تخصیص بصری برای تصویر (Visual attribution)	
جلسه سی و یکم	روش توضیحات خلاف واقع (Counterfactual explanation)	
جلسه سی و دوم	جمع بندی و مرور مطالب	

✓ روش ارزشیابی: تمرین های دستی و کامپیوتری / پروژه برنامه نویسی / امتحان میان ترم / امتحان پایان ترم
 ✓ منابع:

1. Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell. Causal inference in statistics: A primer. John Wiley & Sons, 2016.
2. Shimizu, Shōhei. Statistical causal discovery: LiNGAM approach. Springer, 2022.
3. Jacob, Daniel. "Cate meets ml: Conditional average treatment effect and machine learning." Digital Finance 3.2 (2021): 99-148.
4. Thampi, Ajay. Interpretable AI: Building explainable machine learning systems. Simon and Schuster, 2022.