

خلاصه کتاب درآمدی بر الگوریتم‌های داده‌کاوی و یادگیری ماشین

انتشارات: دانشگاه تهران نویسندگان: دکتر مهدی غضنفری، عاطفه اسمعیل‌پور

مبانی و مفاهیم بنیادین

این اثر به عنوان یک مرجع علمی عمیق، به بررسی پیوند میان ریاضیات بهینه‌سازی و الگوریتم‌های مدرن یادگیری ماشین می‌پردازد. برخلاف متون مقدماتی، این کتاب بر زیرساخت‌های محاسباتی و مدل‌سازی دقیق ریاضی تاکید دارد.

بخش اول: زیربنای ریاضی و بهینه‌سازی

- بهینه‌سازی:** بررسی مفاهیم تحدب، توابع محدب و نقش آن‌ها در همگرایی الگوریتم‌ها. تشریح شرایط کاروش-کوهن-تاکر (KKT) برای مسائل محدودیت‌دار.
- پیچیدگی محاسباتی:** تحلیل زمان و فضای مورد نیاز برای اجرای الگوریتم‌ها و اهمیت بهینگی در داده‌های حجیم.
- توزیع‌های احتمال:** تمرکز بر فرآیندهای گاوسی، شبکه‌های بیزین و مدل‌های مارکوف به عنوان ابزارهای مدل‌سازی عدم قطعیت.

بخش دوم: الگوریتم‌های یادگیری نظارت‌شده

- رگرسیون پیشرفته:** فراتر از رگرسیون خطی، روش‌های تنظیم‌گری مانند Ridge و Lasso برای جلوگیری از بیش‌برازش (Overfitting) مورد بحث قرار می‌گیرند.
- ماشین بردار پشتیبان (SVM):** تحلیل عمیق فضای ویژگی، توابع هسته (Kernel) و یافتن ابرصفحه بهینه.
- مدل‌های گروهی:** بررسی جنگل تصادفی (Random Forest) و روش‌های تقویت (Boosting) برای افزایش دقت پیش‌بینی.

بخش سوم: یادگیری بدون نظارت و کاهش ابعاد

- خوشه‌بندی:** تحلیل الگوریتم‌های k-Means و روش‌های مبتنی بر چگالی مانند DBSCAN با نگاهی به معیارهای ارزیابی درونی و بیرونی.
- تحلیل مولفه‌های اصلی (PCA):** روش‌های استخراج ویژگی و کاهش ابعاد برای تصویرسازی و بهبود کارایی مدل‌ها.

بخش چهارم: یادگیری عمیق و افق‌های نو

- **شبکه‌های عصبی عمیق:** تشریح ساختار لایه‌ها، توابع فعال‌ساز و الگوریتم پس‌انتشار خطا (Backpropagation).
- **معماری‌های مدرن:** نگاهی به شبکه‌های کانولوشنی (CNN) برای داده‌های تصویری و ترنسفورمرها برای پردازش زبان طبیعی.
- **الگوریتم‌های فراابتکاری:** کاربرد الگوریتم‌های ژنتیک و ازدحام ذرات در حل مسائل بهینه‌سازی پیچیده یادگیری ماشین.

جمع‌بندی کاربردی

کتاب با ارائه رویکردی جامع، دانشجو را قادر می‌سازد تا نه تنها از ابزارهای آماده استفاده کند، بلکه منطق پشت هر الگوریتم را درک کرده و در صورت نیاز، آن را برای مسائل خاص بازطراحی نماید.