

# مروری بر مباحث یادگیری ماشین و شناسایی الگو

مناسب برای داوطلبان آزمون دکتری هوش مصنوعی

نویسنده: محمد الیاسی قویی

شابک	0-131-463-600-978:
شماره کتابشناسی	5143679:
ملی	
عنوان و نام پدیدآور	مروری بر مباحث یادگیری ماشین و شناسایی الگو مناسب برای داوطلبان آزمون دکتری هوش مصنوعی/ نویسنده محمد الیاسی قویی .
مشخصات نشر	تهران: شاپرک سرخ، ۱۳۹۷ .
مشخصات ظاهری	۹۷ ص.
موضوع	فراگیری ماشینی -- راهنمای آموزشی (عالی)
موضوع	-- Study and teaching (Higher) Machine learning:
موضوع	فراگیری ماشینی -- آزمون‌ها و تمرین‌ها (عالی)
موضوع	-- Examinations, questions, etc Machine learning: (Higher).
رده بندی دیویی	۳۱/۰۰۶:
رده بندی کنگره	۱۳۹۷ م۴ الف/۳۲۵/QC۵:
سرشناسه	الیاسی قویی، محمد، ۲۰/۰۹/۱۳۶۳، نویسنده
وضعیت فهرست نویسی: فیبا	

عنوان: مروری بر مباحث یادگیری ماشین و شناسایی الگو(مناسب برای داوطلبان آزمون دکتری هوش مصنوعی)

نویسنده: محمد الیاسی قویی

نوبت چاپ: اول بهار ۱۳۹۷

ناشر: شاپرک سرخ

چاپ و صحافی: شاپرک

شمارگان: ۱۰۰۰

قیمت: ۱۶۵۰۰ تومان

شماره شابک: ۰-۱۳۱-۴۶۳-۶۰۰-۹۷۸



## انتشارات شاپرک سرخ

تهران، میدان انقلاب اسلامی، ابتدای خیابان

آزادی، کوچه شهید جنتی، پلاک ۲۴ واحد ۱

۰۹۱۲-۲۱۴۴۶۰۰ ۰۲۱-۶۶۹۳۳۷۴۶

[www.shaparac.com](http://www.shaparac.com)

کلیه حقوق مادی و معنوی اثر متعلق به نویسندگان بوده و هرگونه تکثیر و کپی برداری از

## پیشگفتار

یکی از دغدغه‌های اصلی داوطلبین آزمون دکتری هوش مصنوعی نبود کتاب‌های نکات کلیدی دروس تخصصی ارشد، یعنی یادگیری ماشین و شناسایی الگو است. دروس تخصصی دکتری هوش مصنوعی شامل دو بخش دروس کارشناسی (ساختمان داده و طراحی الگوریتم) است که مجموعاً ۲۰ سوال از این دو درس طرح می‌شود و دروس تخصصی ارشد هم که شامل یادگیری ماشین و شناسایی الگو است مجموعاً ۲۵ سوال از آن‌ها طرح می‌شود. با توجه به اینکه دروس تخصصی ضریب ۴ دارند و دروس عمومی ضریب ۱، موفقیت در دروس تخصصی گام بلندی در دستیابی به رتبه‌های برتر خواهد بود.

کتاب پیش‌رو شامل مباحث اصلی دروس یادگیری ماشین و شناسایی الگو است. بیان مطالب بیشتر به صورت خلاصه و بیان نکات کلیدی است و اصلاً به عنوان یک مرجع پیشنهاد نمی‌شود. اما بعد از مطالعه مراجع اصلی، می‌تواند برای جمع‌بندی مطالب و اطلاع از نکاتی که در کنکور سال‌های قبل آورده شده است مفید باشد. کتاب در چاپ اول قطعاً دارای نواقصی است که انشاءالله در چاپ‌های بعدی برطرف خواهد شد. از عموم مخاطبان این اثر خواهشمند است نظرات ارزشمند خود در زمینه‌ی بهبود محتوای کتاب و نواقص آن را از طریق پست الکترونیک [elyasi.mohamad@gmail.com](mailto:elyasi.mohamad@gmail.com) به اطلاع ما برسانند.

در خاتمه لازم می‌دانم از کمک‌های ارزنده‌ی همسر عزیزم خانم مریم نجاتی نهایت  
تشکر و قدردانی را داشته باشم و این اثر را به ایشان تقدیم می‌دارم.

با آرزوی موفقیت و سربلندی برای همه شما عزیزان

محمد الیاسی قویی

زمستان ۹۶

## فهرست مطالب

- ۱ مروری بر مفاهیم اولیه..... ۹
- ۱-۱ یادگیری ماشین چیست؟..... ۱۰
- ۲-۱ انواع روشهای یادگیری ماشین..... ۱۰
- ۳-۱ تعریف شناسایی الگو..... ۱۲
- ۲ رگرسیون..... ۱۵
- ۱-۲ رگرسیون خطی..... ۱۶
- ۱-۱-۲ یافتن ضرایب رگرسیون با روش کمترین مجموع خطا (SSE)..... ۱۶
- ۲-۲ مدل‌های چند جمله‌ای..... ۱۸
- ۳-۲ وقوع بیش‌برازش (overfitting)..... ۱۸
- ۱-۳-۲ جلوگیری از بروز بیش‌برازش با استفاده از ضریب رگولاسیون..... ۱۹
- ۲-۳-۲ روش نزول در امتداد گرادیان (Gradient descent)..... ۱۹
- ۴-۲ بایاس و واریانس..... ۲۰
- ۳ دسته‌بندی‌های احتمالاتی..... ۲۳
- ۱-۳ مقدمه..... ۲۴
- ۲-۳ روش‌های یادگیری پارامتریک و غیرپارامتریک..... ۲۴

۲۵	..... Discriminative و Generative یادگیری	۳-۳
۲۶	..... قاعده بیز	۴-۳
۲۶	..... MAP روش	۱-۴-۳
۲۷	..... Maximum Likelihood (ML) نمایش درست نمایی	۲-۴-۳
۲۹	..... (Naive Bayes) ساده	۵-۳
۳۰	..... (Bayesian Belief Networks) باور بیزی	۶-۳
۳۴	..... (Gaussian Mixture Models (GMM)) گاوسی	۷-۳
۳۵	..... کمینه کردن میانگین ریسک	۸-۳
۳۹	..... دسته‌بندهای خطی و کاهش ابعاد	۴
۴۰	..... مقدمه	۱-۴
۴۰	..... الگوریتم پرسپترون	۲-۴
۴۱	..... یافتن اوزان بهینه	۱-۲-۴
۴۲	..... تحلیل جداساز خطی یا جداساز خطی فیشر (LDA)	۳-۴
۴۵	..... تحلیل مؤلفه‌های اصلی (PCA)	۴-۴
۴۹	..... ماشین بردار پشتیبان (SVM)	۵-۴
۵۲	..... کرنل در SVM چیست؟	۱-۵-۴

۵۳	.....SVM در حاشیه نرم در ۲-۵-۴
۵۹	..... ترکیب دسته‌بندی کننده‌ها ۵
۶۰	..... مقدمه ۱-۵
۶۰	..... تفاوت ۲-۵ Ensemble Learning با Multiple classifier
۶۱	..... روش ۳-۵ Bagging
۶۲	..... روش ۴-۵ Boosting
۶۲	..... الگوریتم ۱-۴-۵ Adaboost
۶۵	..... درخت تصمیم ۶
۶۶	..... یادگیری مبتنی بر درخت تصمیم ۱-۶
۶۷	..... الگوریتم ۲-۶ ID3 برای یادگیری درخت تصمیم
۷۱	..... تعریف بیش برآزش ۳-۶
۷۲	..... بیش برآزش در درخت تصمیم ۱-۳-۶
۷۷	..... یادگیری مبتنی بر نمونه‌ها ۷
۷۸	..... مقدمه ۱-۷
۷۹	..... نزدیکترین همسایه K-۲-۷
۸۰	..... نمودار ۱-۲-۷ Voronoi

- ۳-۷ ویژگی‌های یادگیری مبتنی بر نمونه ..... ۸۲
- ۴-۷ خوشه‌بندی ..... ۸۴
- ۱-۴-۷ خوشه‌بندی سلسله‌مراتبی ..... ۸۴
- ۲-۴-۷ خوشه‌بندی طیفی ..... ۸۵
- ۵-۷ الگوریتم K-means ..... ۸۶
- ۸ مباحث متفرقه ..... ۸۹**
- ۱-۸ بعد VC یک فرضیه ..... ۹۰
- ۲-۸ تخمین خطای واقعی یک دسته‌بند ..... ۹۳