

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

آموزش جامع نرم افزار

WEKA

(یادگیری ماشین)

مؤلفان:

رمکو بویبیکارت - فرانک ایبن - ریچارد کیرکبی - پیتر روتمن
الکس سیوالد - دیوید جی. جی. جی.

مترجمان:

دکتر جابر کریم پور

حامد بابایی

حمید جبرئیلی

عنوان و نام پدیدآور	آموزش جامع نرم افزار WEKA (یادگیری ماشین) // مولفان رمکو بوویبیکارت... ا. و دیگران؛ مترجمان جابر کریم پور، حامد بابایی، حمید چیرئیلی
مشخصات نشر	تیریز؛ انتشارات پژوهش های دانشگاه، ۱۳۹۵
مشخصات ظاهری	۳۸۰ ص. مصور شمردار
شابک	۹۷۸-۶۰۰-۸۵۵۵-۶۷-۴
وضعیت فهرست نویسی	فیبیا
موضوع	فرآگیری ماشینی -- نرم افزار
موضوع	Machine learning -- Computer software
موضوع	برنامه نویسی
موضوع	Computer programming
شناسه افزوده	بوکرت، رمکو آر.
شناسه افزوده	Bouckaert, Remco R
شناسه افزوده	کریم پور، ج. / ۱۳۵۲ -
شناسه افزوده	بابایی، حامد، ۱۳۶۳ -
شناسه افزوده	چیرئیلی، حمید، ۱۳۶۱ -
شناسه افزوده	انتشارات پژوهش های دانشگاه --
رده بندی کنگره	۱۳۹۵ / QC۱۱
رده بندی دیویی	۰۰۶/۳۱
شماره کتابشناسی ملی	۴۴۹۵۸۲۷

عنوان : آموزش جامع نرم افزار WEKA (یادگیری ماشین)	مؤلف : جابر کریم پور، حامد بابایی، حمید چیرئیلی
شابک : ۹۷۸-۶۰۰-۸۵۵۵-۶۷-۴	قیمت : ۱۵۰۰۰ ریال
صفحه آرای : محمد محمد خانی	طراح روی جلد : جابر بابایی
چاپ و صحافی : مدیران	قطع : وزیری

انتشارات پژوهش های دانشگاه



انتشارات پژوهش های دانشگاه
پیرا، استاد، استاد، استاد، استاد

مرکز تخصصی چاپ نشر کتب دانشگاه و نهادهای علمی هیات علم دانشگاهها

آدرس دفتر مرکزی : آذربایجان شرقی، تبریز، آبرسان، مجتمع تجاری بلور، رویروی کلانتری، داخل حیاط (ضلع شرقی)، پلاک ۳۹

مرکز بخش : آذربایجان شرقی، تبریز، خیابان امام، رویروی سه راه طالقانی، انتشارات شایسته / ۳۵۵۶۱۹۶۱-۴۱

تلفن های تماس بخش مدیریت : ۰۹۱۴۶۱۲۳۳۶۶ / ۰۴۱۳۲۲۶۲۱۱۶ / بخش امور فنی : ۰۹۱۴۵۰۱۹۸۵۰ / ۰۴۱۳۲۲۵۰۲۴۸

www.Unpb.org / Info@unpb.org

www.Pajohesh-daneshgah.ir / Info@pajohesh-daneshgah.ir

طبق قانون حمایت از حقوق ناشران و مولفان هر شخص حقیقی و حقوقی که تمام یا قسمتی از این اثر را بدون اجازه ناشر و مولف، نشر یا عرضه یا تکثیر یا تجدید چاپ نماید مورد پیگیری جدی قانونی خواهد گرفت و مطابق با جرایم قوانین اسلامی بر خورد خواهد شد.

فهرست مطالب

۱۳	پیشگفتار
فصل اول Command-line premier	
۱۴	۱-۱ مقدمه
۱۴	۲-۱ مفاهیم اولیه
۱۴	۱-۲-۱ مجموعه داده
۱۶	۱-۲-۲ Classifier
۱۶	۱-۲-۳ weka.filters
۲۰	۱-۲-۴ weka.classifier
۲۷	۳-۱ مثالها
شروع کار با نرم افزار weka	
فصل سوم Simple GUI	
۳۶	۱-۳ فرمانها
۳۷	۲-۳ Invocation
۳۸	۳-۳ command redirection
۳۸	۴-۳ Command completion
فصل چهارم Explorer	
۴۲	۱-۴ واسطه کاربر
۴۲	۱-۱-۴ explorer weka سرببرگهای پنجره
۴۳	۲-۱-۴ Status Box
۴۳	۳-۱-۴ Log Button
۴۴	۴-۱-۴ WEKA Status Icon
۴۴	۵-۱-۴ graphical output
۴۵	۲-۴ Preprocessing
۴۵	۱-۲-۴ Loading data
۴۶	۲-۲-۴ The Current Relation
۴۶	۳-۲-۴ Working With Attributes

۴۸	Working With Filters	۴-۲-۴
۵۱	Classification	۲-۴
۵۱	انتخاب classifier	۱-۲-۴
۵۲	Test Option	۲-۲-۴
۵۴	The Class Attribute	۲-۲-۴
۵۵	Training a Classifier	۴-۲-۴
۵۵	The Classifier Output Text	۵-۲-۴
۵۶	The Result List	۶-۲-۴
۵۸	Clustering	۴-۴
۵۸	clustreer انتخاب یب	۱۱-۴-۴
۵۸	Cluster Model	۲-۴-۴
۵۹	Ignoring Attributes	۲-۴-۴
۵۹	کار با فیلترها	۴-۴-۴
۵۹	یادگیری clusterها	۵-۴-۴
۶۰	Associating	۵-۴
۶۰	Setting Up	۱-۵-۴
۶۱	Association	۲-۵-۴
۶۱	Selecting Attributes	۶-۴
۶۱	Searching and Evaluating	۱-۶-۴
۶۲	انتخابها	۲-۶-۴
۶۲	selection اجرای	۳-۶-۴
۶۴	Visualizing	۷-۴
۶۴	۱-۷-۴ ماتریس ترسیم پراکندگی	
۶۴	۲-۷-۴ انتخاب و ترسیم پراکندگی دو بعدی مستقل	
۶۵	۳-۷-۴ انتخاب نمونه ها	

فصل پنجم Experimenter

۶۸	۱-۵ مقدمه	
۶۹	Standard Experiments	۲-۵
۶۹	Simple	۱-۲-۵
۸۳	Advanced	۲-۲-۵
۱۰۸	۳-۵ آزمایش های خوشه	
۱۱۲	۴-۵ آزمایش های راه دور	
۱۱۲	۱-۴-۵ آمادسازی	

۱۱۲	server تنظیمات پایگاه داده	۲-۴-۵
۱۱۳	تنظیمات ماشین راه دور	۳-۴-۵
۱۱۴	Experiments پیکربندی	۴-۴-۵
۱۱۶	Multi-core support	۵-۴-۵
۱۱۶	Troubleshooting	۶-۴-۵
۱۱۸	تحلیل نتایج	۵-۵-۵
۱۱۸	تنظیمات	۱-۵-۵
۱۲۲	ذخیره‌ی نتایج	۲-۵-۵
۱۲۳	تغییر الگوریتم مبنا	۳-۵-۵
۱۲۴	معنی آمارها	۴-۵-۵
۱۲۴	خلاصه‌ی آمارها	۵-۵-۵
۱۲۵	Ranking Test	۶-۵-۵

فصل ششم KnowledgeFlow

۱۲۸	مقدمه	۱-۶
۱۲۹	ویژگی‌ها	۲-۶
۱۳۰	ابزارها	۳-۶
۱۳۰	DataSources	۱-۳-۶
۱۳۰	DataSinks	۲-۳-۶
۱۳۰	Filters	۳-۳-۶
۱۳۰	Classifiers	۴-۳-۶
۱۳۱	Clusterers	۵-۳-۶
۱۳۱	Evaluation	۶-۳-۶
۱۳۳	نمونه	۴-۶
۱۳۳	Cross-validated J48	۱-۴-۶
۱۳۵	Plotting multiple ROC curves	۲-۴-۶
۱۳۹	Processing data incrementally	۳-۴-۶
۱۴۱	Plugin Facility	۵-۶

فصل هفتم ArffViewer

۱۴۵	Menus	۱-۷
۱۴۷	Editing	۲-۷

فصل هشتم Bayesian Network Classifiers

۱۵۰	مقدمه	۱-۸
۱۵۵	Local score based structure learning	۲-۸

۱۵۶	Local score metrics	۱-۲-۸
۱۵۸	الگوریتمهای جستجو	۲-۲-۸
۱۶۱	Conditional independence test based structure learning	۳-۸
۱۶۴	Global score metric based structure learning	۴-۸
۱۶۵	Fixed structure 'learning'	۵-۸
۱۶۶	Distribution learning	۶-۸
۱۶۷	Running from the command line	۷-۸
۱۷۸	Inspecting Bayesian networks	۸-۸
۱۸۳	Bayes Network GUI	۹-۸
۱۹۸	Bayesian nets in the experiment	۱۰-۸
۱۹۸	Adding your own Bayesian network learners	۱۱-۸
۲۰۱	FAQ	۱۲-۸
۲۰۳	Future development	۱۳-۸
	فصل نهم ARFF	
۲۰۶	۱- مرور کلی	۹
۲۰۷	Examples	۲-۹
۲۰۷	The ARFF Header Section	۱-۲-۹
۲۱۰	The ARFF Data Section	۲-۲-۹
۲۱۱	Sparse ARFF files	۳-۹
۲۱۲	Instance weights in ARFF files	۴-۹
	فصل دهم XRFF	
۲۱۶	File extensions	۱-۱۰
۲۱۶	Comparison	۲-۱۰
۲۱۶	ARFF	۱-۲-۱۰
۲۱۶	XRFF	۲-۲-۱۰
۲۱۸	Sparse format	۳-۱۰
۲۱۹	Compression	۴-۱۰
۲۲۰	Useful features	۵-۱۰
۲۲۰	Class attribute specification	۱-۵-۱۰
۲۲۰	Attribute weights	۲-۵-۱۰
۲۲۰	Instance weights	۳-۵-۱۰
	فصل یازدهم Converters	
۲۲۴	۱-۱ مقدمه	۱۱

۲۲۴		Usage۲-۱۱
۲۲۴		File converters۱-۲-۱۱
۲۲۵		Database converters۲-۲-۱۱

فصل دوازدهم Stemmers

۲۲۸		۱-مقدمه ۱۲
۲۲۸		Snowball stemmers۲-۱۲
۲۲۹		Using stemmers۳-۱۲
۲۲۹		Commandline۱-۳-۱۲
۲۲۹		StringToWordVector۲-۳-۱۲
۲۲۹		Adding new stemmers۴-۱۲

فصل سیزدهم Databases

۲۳۲		Configuration files۱-۱۳
۲۳۳		Setup۲-۱۳
۲۳۴		Missing Datatypes۳-۱۳
۲۳۵		Stored Procedures۴-۱۳
۲۳۶		Troubleshooting۵-۱۳

فصل چهاردهم Windows database

فصل پانزدهم Research

۲۴۶		Citing Weka۱-۱۵
۲۴۶		Paper references۲-۱۵

فصل شانزدهم Using the API

۲۵۳		Option handling۱-۱۶
۲۵۶		Loading data۲-۱۶
۲۵۷		Loading data from files۱-۲-۱۶
۲۵۸		Loading data from databases۲-۲-۱۶
۲۶۰		۱۶-۳ ایجاد مجموعه‌ی داده‌ها در حافظه
۲۶۱		۱۶-۳-۱ تعریف قالب
۲۶۳		۱۶-۳-۲ اضافه کردن داده
۲۶۵		۱۶-۴ تصادفی کردن داده
۲۶۶		۱۶-۵ فیلترسازی
۲۶۸		۱۶-۵-۱ فیلترسازی دسته‌ای
۲۶۸		۱۶-۵-۲ on-the-fly فیلترسازی
۲۷۰		۱۶-۶ طبقه‌بندی

۲۷۰	۱۶-۶-۱ ساخت یک طبقه‌بندی کننده
۲۷۲	۱۶-۶-۲ ارزیابی یک طبقه‌بندی کننده
۲۷۶	۱۶-۶-۳ طبقه‌بندی نمونه‌ها
۲۷۸	۱۶-۷-۱ خوشه‌بندی
۲۷۸	۱۶-۷-۱ ساخت یک خوشه‌ساز
۲۸۰	۱۶-۷-۲ ارزیابی یک خوشه‌ساز
۲۸۲	۱۶-۷-۳ خوشه سازی نمونه ها
۲۸۳	۱۶-۸ انتخاب صفات
۲۸۵	۱۶-۸-۱ استفاده از ابرطبقه‌بندی کننده‌ها
۲۸۶	۱۶-۸-۲ استفاده از فیلتر
۲۸۷	۱۶-۱۳ مفادهی مستقیم از API
۲۸۸	۱۶-۹ ذخیره‌ی داده
۲۸۸	۱۶-۹-۱ ذخیره داده در فایلها
۲۸۹	۱۶-۹-۲ ذخیره‌ی داده در پایگاه داده
۲۹۰	۱۶-۱۰ تجسم
۲۹۱	۱۶-۱۰-۱ منحنیهای ROC
۲۹۲	۱۶-۱۰-۲ گرافها
۲۹۴	۱۶-۱۱ اتسلسل (توالی)

فصل هفدهم توسعه WEKA

۲۹۸	۱۷-۱ نوشتن یک طبقه بندی کنندهی جدید
۲۹۸	۱۷-۱-۱ انتخاب کلاس پایهای
۳۰۰	۱۷-۱-۲ واسطه‌های افزوده
۳۰۰	۱۷-۱-۳ بسته‌ها
۳۰۱	۱۷-۱-۴ پیاده‌سازی
۳۱۴	۱۷-۲ نوشتن یک فیلتر جدید
۳۱۵	۱۷-۲-۱ Default approach
۳۲۴	۱۷-۲-۲ Simple approach
۳۲۹	۱۷-۲-۳ Capabilities
۳۲۹	۱۷-۲-۴ Packages
۳۳۰	۱۷-۲-۵ Revisions
۳۳۰	۱۷-۲-۶ Testing
۳۳۱	۱۷-۳ Writing other algorithms
۳۳۲	۱۷-۳-۱ Clusterers

۳۳۴	Attribute selection	۲-۳-۱۷
۳۳۶	Associators	۳-۳-۱۷
۳۳۸	Extending the Explorer	۴-۱۷
۳۳۸	Adding tabs	۱-۴-۱۷
۳۴۹	Adding visualization plugins	۲-۴-۱۷

فصل هجدهم مستندات تکنیکی

۳۵۴	ANT	۱-۱۸
۳۵۴	مبانی	۱-۱-۱۸
۳۵۴	ANT و weka	۲-۱-۱۸
۳۵۵	CLASSPATH	۲-۱۸
۳۵۵	CLASSPATH تنظیمات	۱-۲-۱۸
۳۵۶	RunWeka.bat	۲-۲-۱۸
۳۵۷	java-jar	۳-۲-۱۸
۳۵۸	Subversion	۳-۱۸
۳۵۸	کلیات	۱-۳-۱۸
۳۵۸	کد منبع	۱۸.۳.۲
۳۵۸	JUnit	۳-۳-۱۸
۳۵۹	نسخه مخصوص	۴-۳-۱۸
۳۵۹	Clients	۵-۳-۱۸
۳۶۰	GenericObjectEditor	۴-۱۸
۳۶۰	مقدمه	۱-۴-۱۸
۳۶۱	ساختار فایل	۲-۴-۱۸
۳۶۲	Exclusion	۳-۴-۱۸
۳۶۳	Class Discovery	۴-۴-۱۸
۳۶۳	سلسله مراتب کلاس های چندگانه	۵-۴-۱۸
۳۶۴	قابلیتها	۶-۴-۱۸
۳۶۵	خصوصیات	۵-۱۸
۳۶۶	اولویت	۱-۵-۱۸
۳۶۶	مثالها	۲-۵-۱۸
۳۶۶	XML	۶-۱۸
۳۶۷	خط فرمان	۱-۶-۱۸

Other resources فصل نوزدهم

۳۷۲	Mailing list	۱-۱۹
-----	-------------------	------

۳۷۲	 Troubleshooting ۲-۱۹
۳۷۲	 weka دانلود ۱-۲-۱۹
۳۷۲	 OutOfMemoryException ۲-۲-۱۹
۳۷۳	 Mac OSX ۳-۲-۱۹
۳۷۳	 StackOverflowError ۴-۲-۱۹
۳۷۴	 just-in-time (JIT) compiler ۵-۲-۱۹
۳۷۴	 CSV file conversion ۶-۲-۱۹
۳۷۴	 فایل ARFF بار نمی شود ۷-۲-۱۹
۳۷۴	 ARFF فضای خالی در برجسپ فایلهای ۸-۲-۱۹
۳۷۴	 CLASSPATH ۹-۲-۱۹
۳۷۵	 ID نمونه ۱۰-۲-۱۹
۳۷۵	 Visualization ۱۱-۲-۱۹
۳۷۶	 ۱۲-۲-۱۹ مصد حافظه و زیاده روب
۳۷۶	 GUIChooser starts but not Experiment Explorer ۱۳-۲-۱۹
۳۷۶	 ابزارهای KnowledgeFlow خالی هستند ۱۴-۲-۱۹
۳۷۷	 لینکها ۱۵-۲-۱۹
۳۷۸	 Bibliography

پیشگفتار

ما در عصر اطلاعات و ارتباطات به سر می‌بریم. در این دوران، کامپیوتر و فناوری اطلاعات و ارتباطات، نقش عمده‌ای در رشد اقتصادی، اجتماعی، فرهنگی، و سیاسی یک جامعه دارد. پیشرفت در زمینه‌های مختلف علوم آرزوی همیشگی بشریت می‌باشد که علم کامپیوتر یکی از مولفه‌ها و بال‌های پرواز این حرکت همیشه پویا و رو به جلو برای تسریع، دقت، انضباط و بهینگی خواهد بود. در این راستا حوزه‌ی هوش مصنوعی که دگرگونی‌های شگرفی در کلیه علوم مختلف علی‌الخصوص علم کامپیوتر داشته است. یکی از زیر شاخه‌های اساسی هوش مصنوعی سیستم‌های خبره (یادگیری ماشین) می‌باشد که این حوزه برای تجزیه و تحلیل و حل بسیاری از مسائل ساده تا پیچیده هیافته‌ای نرین و بهینه ارائه نموده است. کتاب حاضر حاصل ترجمه نرم افزار weka که یک نرم افزار کاربردی برای شبیه‌سازی، آزمایش و ارائه عملی از به کارگیری رهیافته‌های یادگیری ماشین برای حل مسائل متنوع می‌شد. در این کتاب سعی بر آن شده که کلیه المان‌های نرم افزار weka به طور کامل توضیح داده شوند و امانت‌داری نویسندگان اصلی نرم افزار حفظ گردد. کتاب حاضر مشتمل بر ۱۹ فصل می‌باشد که به صورت گام به گام فرآیند مطالعه‌ی هر فصل با رعایت همپوشانی فصول با یکدیگر انجام شده است.

در پایان شایسته است از همه‌ی عزیزانی که در تهیه‌ی این کتاب همکاری ارزنده‌ای داشته‌اند، سپاسگزاری و قدردانی شود. همچنین از تمامی خوانندگان گرمای خواهشمندیم نظرات، پیشنهادات و انتقادات خود را از طریق ایمیل در اختیار مؤلفان یا ناشر قرار دهند تا در چاپ‌های بعدی مورد توجه و استفاده قرار گیرد.

جابر کریم‌پور

Karimpour@tabrizu.ac.ir

حامد بابایی

Hamedbabaei۶۳@gmail.com

حمید جبرئیلی

Hamid.jabraili@gmail.com