

শূন্য থেকে পাইথন মেশিন লার্নিং

হাতেকলমে সাইকিট-লার্ন

রকিবুল হাসান

হাতেকলমে মেশিন লার্নিং সিরিজ
আইরিস ডেটাসেট প্রজেক্ট





প্রকাশক : আদর্শ

৩৮ পি কে রায় রোড, বাংলাবাজার (২য় তলা), ঢাকা ১১০০

কনকর্ড এম্পোরিয়াম, কাঁটাবন, ঢাকা ১২০৫

☎+০২-৯৬১২৮৭৭, ০১৭৯৩২৯৬২০২, ০১৭১০৭৭৯০৫০

info@adarshapublications.com

www.adarshapublications.com

শূন্য থেকে পাইথন মেশিন লার্নিং : হাতেকলমে সাইকিট-লার্ন

১ম প্রকাশ : ১৭ আষাঢ় ১৪২৬; ১ জুলাই ২০১৯

© রকিবুল হাসান

সর্বস্বত্ত্ব সংরক্ষিত; লেখক, প্রকাশক ও অনুবাদকের লিখিত অনুমতি ব্যতীত
বইটি বা বইটির অংশবিশেষ যেকোনো মাধ্যমে প্রকাশ সম্পূর্ণ নিষিদ্ধ

প্রচ্ছদ : মোস্তাফিজ কারিগর

মুদ্রণ ব্যবস্থাপনা : আদর্শ প্রিন্টার্স

রকমারিতে আদর্শের বই : www.rokomari.com/adarsha

মূল্য : বাংলাদেশে ৩০০ টাকা

*Shunno Theke Python Machine Learning:
Hate Kalame Scikit Learn (Published in Bengali)
by Rakibul Hassan*

Published by Adarsha

38 P. K. Ray Road, Banglabazar (1st floor), Dhaka 1100
Concord Emporium, Kataban, Dhaka 1205

ISBN: 978-984-8040-38-6

সূচি

কৃতজ্ঞতা	১১
মুখবন্ধ	১৩
বইটা কীভাবে কাজে লাগাবেন?	১৫
লেখার পেছনের দর্শন	১৭
কেন বইটা লিখতে চাইলাম?	১৭
কীভাবে শিখব?	১৯
শেখার পেছনের দর্শন	২১
স্ট্যান্ডিং অন দ্য সোল্ডারস অব জায়ান্টস	২৩
মেশিন লার্নিং কী?	২৬
মেশিন লার্নিং জিনিসটা কী?	২৬
টেক্সটবুকের মেশিন লার্নিং	২৮
অঙ্কে মেশিন লার্নিং ১	২৯
অঙ্কে মেশিন লার্নিং ২, বিঁঝি পোকা এবং তাপমাত্রা	৩২
অঙ্কে মেশিন লার্নিং ৩, লিনিয়ার রিগ্রেশন, সরলরেখার ইকুয়েশন	৩৭
অঙ্কে মেশিন লার্নিং ৪, লিনিয়ার রিগ্রেশন, প্রথম প্রেডিকশন	৪১
কস্ট ফাংশন, লস অথবা এরর	৪৬
অঙ্কে মেশিন লার্নিং ৫, অ্যাক্যুরেসি, লিস্ট স্কয়ার রিগ্রেশন	৪৭

দরকারি কিছু টুলস	৫২
এক প্যাকেজে সবকিছু : জুপিটার ও গুগল কোলাব নোটবুক	৫২
সাইকিট-লার্ন	৫৪
পান্ডাজ	৫৫
নামপাই	৫৭
ডেটা সায়েন্সে পাইথন	৫৮
মেশিন লার্নিং থার্ড পার্টি লাইব্রেরিতে ব্যবহার করা বিভিন্ন মডিউল	৬২
জুপিটার নোটবুকের খুঁটিনাটি	৬৩
সুপারভাইজড লার্নিং	৬৬
মডেলের জেনারেলাইজেশন, ওভার-ফিটিং এবং আন্ডার-ফিটিং	৬৯
আইরিস ডেটাসেট	৭২
কেন এই ডেটাসেট?	৭২
সাইকিট-লার্নে আইরিস ডেটাসেট	৭৪
সাইকিট-লার্নের ডেটা লে-আউট, ডেটা হ্যান্ডলিং	৭৭
সাইকিট-লার্ন এর ডেটা নিয়ে কাজ করার ধারণা	৮০
কিছু মেশিন লার্নিং টার্মিনোলজি	৮২
মেশিন লার্নিং মডেলের জন্য কী দরকার?	৮৪
ডেটা ভিজুয়ালাইজেশন	৮৫
এক্সপ্লোরেরি ডেটা অ্যানালাইসিস	৯২
সাইকিট-লার্ন-এ ডেটা হ্যান্ডলিংয়ের নিয়ম	৯৫
মডেল ইভালুয়েশনের ধারণা	৯৬
সাইকিট-লার্ন ‘এস্টিমেটর’	৯৮
এস্টিমেটরের কাজের ধাপ (কে নিয়ারেস্ট নেইবার ক্লাসিফায়ার দিয়ে)	৯৯
কোডে প্রথম মডেল ও প্রেডিকশন (কেএনএন দিয়ে)	১০৪
কোডে প্রথম মডেল ও প্রেডিকশন (লজিস্টিক রিগ্রেশন দিয়ে)	১০৬
মডেলের ইভালুয়েশন, দুটো অ্যালগরিদম	১০৮
ট্রেনিং ডেটার ওপর ইভালুয়েশন	১০৮

ট্রেনিং ও টেস্ট ডেটা ভাগ করে ইভালুয়েশন	১১২
নেইবারের সংখ্যা কত হলে মডেলের অ্যাক্যুরেসি ভালো?	১১৭
মডেলের মধ্যে কমপ্লেক্সিটি আর জেনেরালাইজেশনের সম্পর্ক	১১৯
ক্রস ভ্যালিডেশনের প্যারামিটারের টিউনিং, মডেল সিলেকশন	১২১
অ্যাক্যুরেসি বাড়তে আইরিস ডেটাসেটের জন্য প্যারামিটার টিউনিং	১২৫
ছোট্ট একটা লিনিয়ার ক্লাসিফিকেশন (ফিচার স্কেলিংসহ)	১২৭
ডিসিশন ট্রি কীভাবে কাজ করে? খালি চোখে আইরিস ডেটাসেট	১৩৪
খালি চোখে প্রেডিকশন, ট্রেনিং ইনস্ট্যান্স	১৩৯
ডাইমেনশনালিটি রিডাকশন, ফিচার সিলেকশন, ফিচার ইম্পর্ট্যান্স	১৪১

মেশিন লার্নিং ভবিষ্যৎ ধারণা	১৪৩
কী করব সামনে?	১৪৩
আরও দূরে, বহুদূরে—অসাধারণ কিছু বই	১৪৫
কীভাবে রেকমেন্ডার সিস্টেম কাজ করে?	১৪৬
বাংলাদেশ ও সামনের পাঁচ বছর	১৫০
পরবর্তী সাহায্য, লেখকের সঙ্গে যোগাযোগ	১৫২