

「機械学習を使いたい人のための入門講座」 担当

- 比戸将平 (HIDO Shohei)
- TwitterID: @sla
- 専門：データマイニング、機械学習
- 経歴：
 - 2006-2012: IBM東京基礎研究所データ解析グループ
 - 機械学習のアルゴリズム研究開発（主に異常検知）
 - +いわゆるデータサイエンティスト的なお客様仕事
 - 2012-: 株式会社Preferred Infrastructure
 - 大規模オンライン分散機械学習基盤Jubatusチームリーダー
 - 2014-: 株式会社Preferred Networks
 - 2015-: Preferred Networks America, Inc. @ シリコンバレー
 - Chief Research Officer



2014/08/26 Machine Learning Casual Talks #2

あなたの仕事に 機械学習を活かす 5つのポイント

株式会社Preferred Infrastructure

比戸 将平

Clip slide

そう、「機械学習は苦しい」！！！！

- 機械学習を使えば楽しく問題が解決する？
- 機械学習を使えば楽しくお金が儲かる？
- 機械学習を使えば誰でも有用な知見が得られる？
- 機械学習を使えば誰でも素晴らしいサービスを作れる？
- 機械学習を覚えればワンランク上の技術者になれる？
- 機械学習を覚えれば社内でも一目置かれる人になれる？
- 機械学習を覚えればきっと新しい自分になれる？
- 機械学習を覚えれば明日から自分もモテモテになる？

2014/01/17 エンジニアサポートCROSS 2014 機械学習CROSS セッション主催



2013/08/22 PFI公開セミナー データサイエンティスト三部作：3作目

Clip slide

PFIセミナー 2013/08/22

さらば！
データサイエンティスト

株式会社Preferred Infrastructure
リサーチャー&Jubatusチームリーダー
比戸 将平

2013/03/28 PFI公開セミナー データサイエンティスト三部作：2作目

Clip slide

PFIセミナー 2013/03/28

(道具としての) データサイエンティストの つかい方

株式会社Preferred Infrastructure
リサーチャー&Jubatusチームリーダー
比戸 将平

2012/09/13 PFI公開セミナー データサイエンティスト三部作：1作目

Clip slide

PFIセミナー 2012/09/13

データサイエンティストの つくり方

株式会社Preferred Infrastructure
リサーチャー & Jubatusチームリーダー
比戸 将平

データサイエンティストのお仕事

1. 業務でデータを使って改善できそうなところを探す
2. 改善のための仮説を立てて提案を作る
3. 社内で調整してプロジェクト化する
4. データを収集する・データを受け取る
5. データをクリーニングする
6. データを可視化する
7. データを集計する
8. データマイニングや機械学習を用いる
9. 結果を元に再調整や再実験を繰り返す
10. 結果を可視化してレポートにまとめる
11. レポートを関係者と共有する
12. 結果を元に有効と思われるビジネスアクションを取る
13. ビジネスアクションの効果を測定して検証する
14. (可能であればこれらをPDCAサイクルのように回す)



うおおおおお

データサイエンティスト

養成読本で伝えたかったこと： Demystifying Machine Learning

- つまり「機械学習の化けの皮を剥がす」
- 神話 1
 - 「数学も統計もわからないし、難しくて自分にはとても使えない、使いこなせない」
- 神話 2
 - 「よくわからないけど機械学習の専門家に今データで困っていることは全部解決できそうだ」