

Deep Learningによるビッグデータ解析 ～手法やCUDAによる高速化

2014年7月16日

G-DEPソリューションパートナー
株式会社 システム計画研究所
奥村 義和

目次



- ・DeepLearningとGPU
- ・G-DEPテストドライブ
- ・ビッグデータ・GPU・DeepLearningの接点

目次



- ・DeepLearningとGPU
 - ・DeepLearningとは
 - ・仕組みと計算
 - ・初期化の1手法：AutoEncoderとその応用
- ・G-DEPテストドライブ
- ・ビッグデータ・GPU・DeepLearningの接点

DeepLearningとは



- ・多層NNを使った機械学習
- ・認識問題で高い性能

- ・一般物体認識

74% → 85%

- ・自然言語処理
- ・音声認識

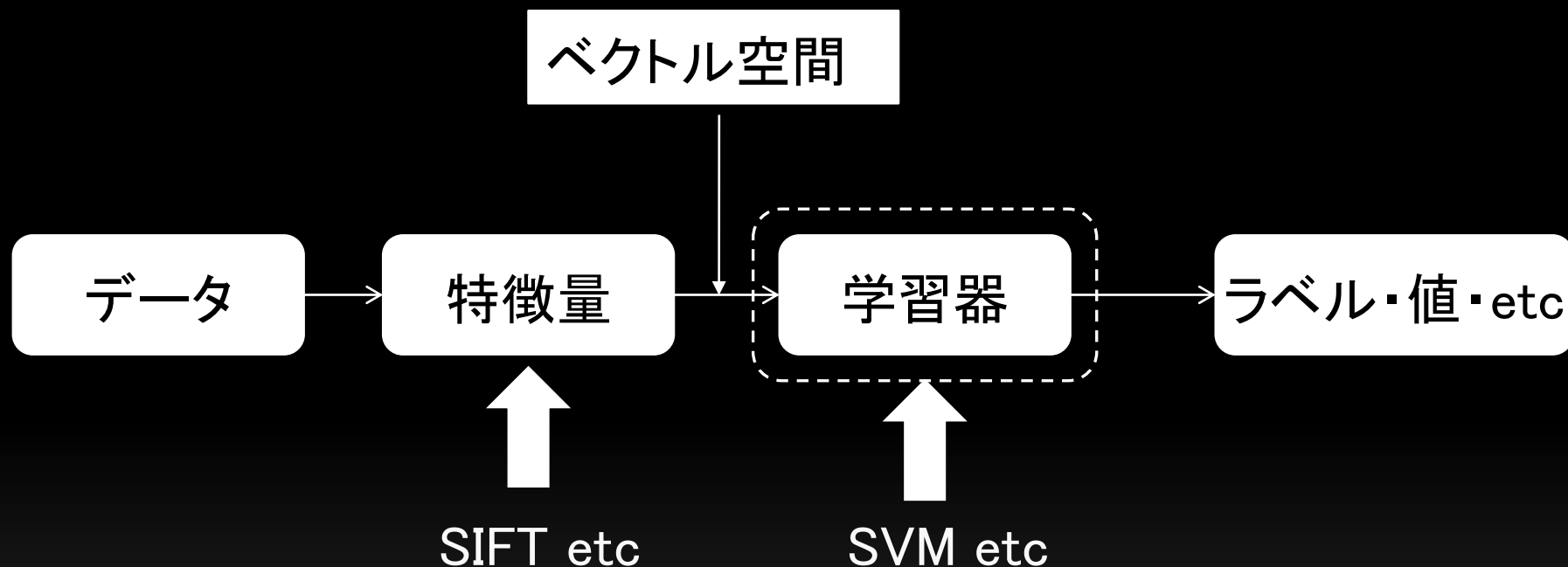
- ・その他

- ・顔認証 人間に肉薄！
- ・デノイジング より自然に！

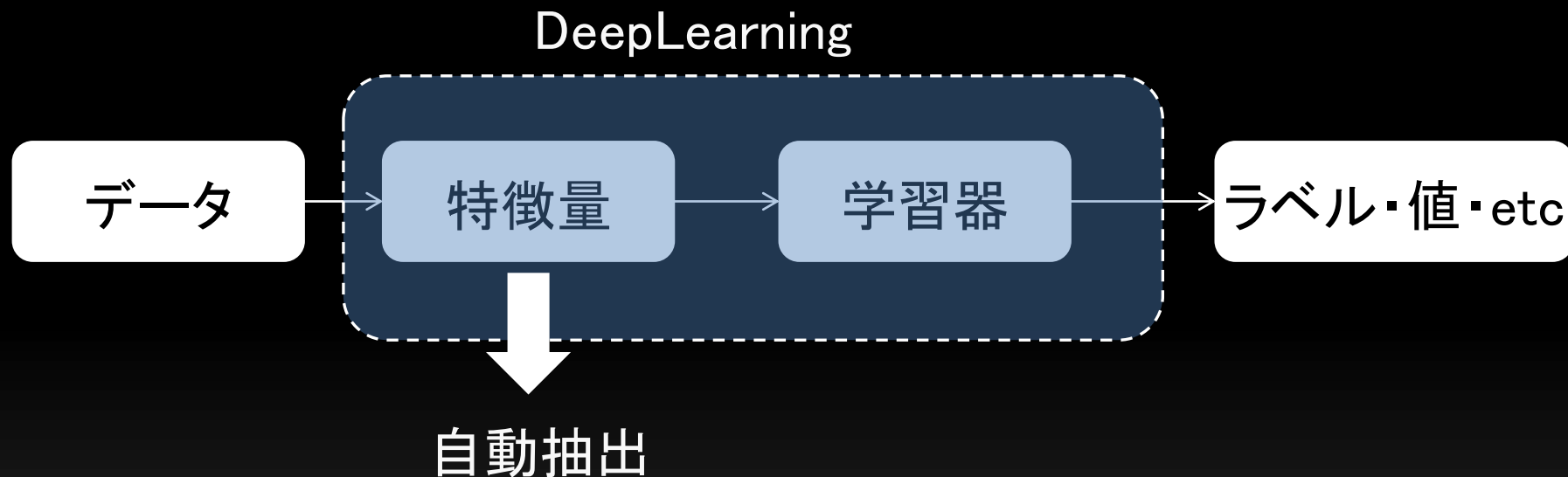


→ cherry

認識問題への従来アプローチ

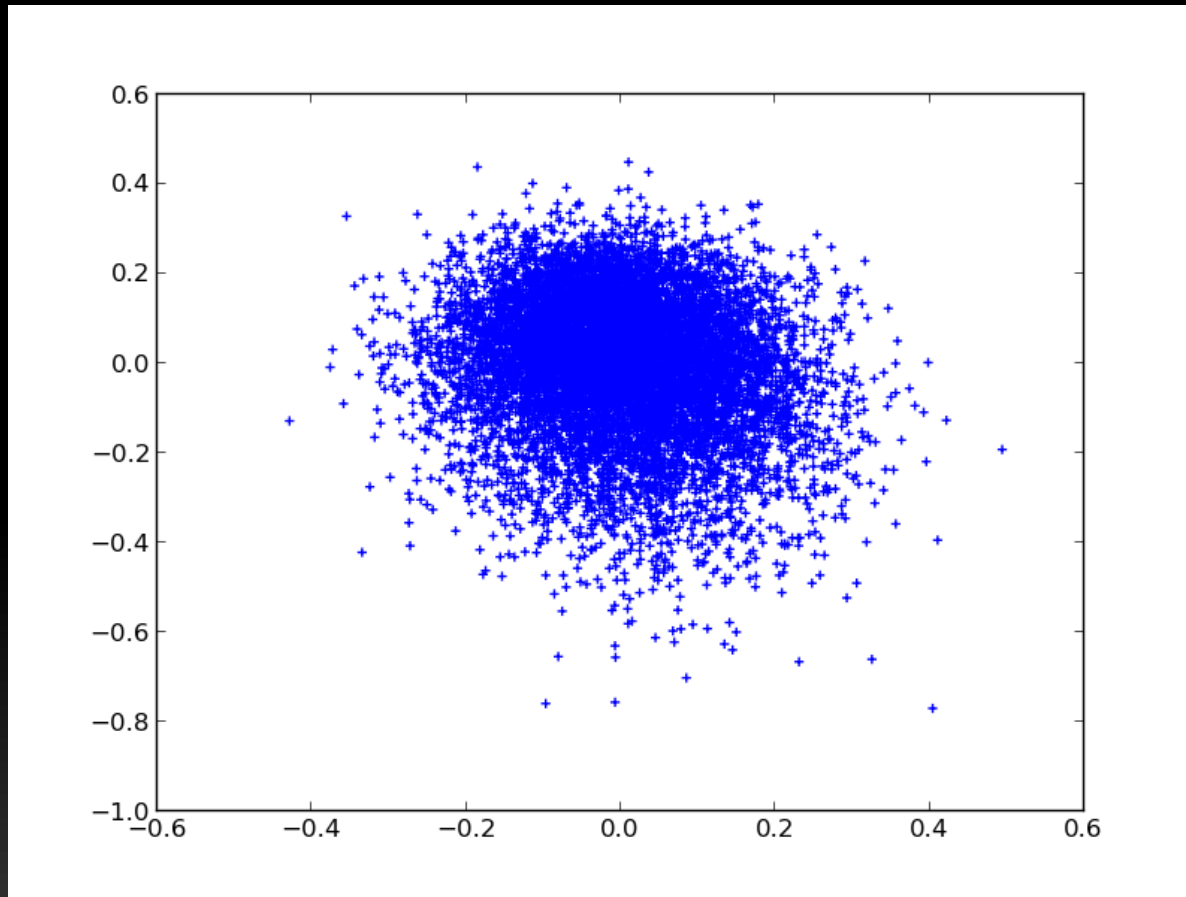


- 特徴量は研究者が定義
- 問題領域毎に解が異なる



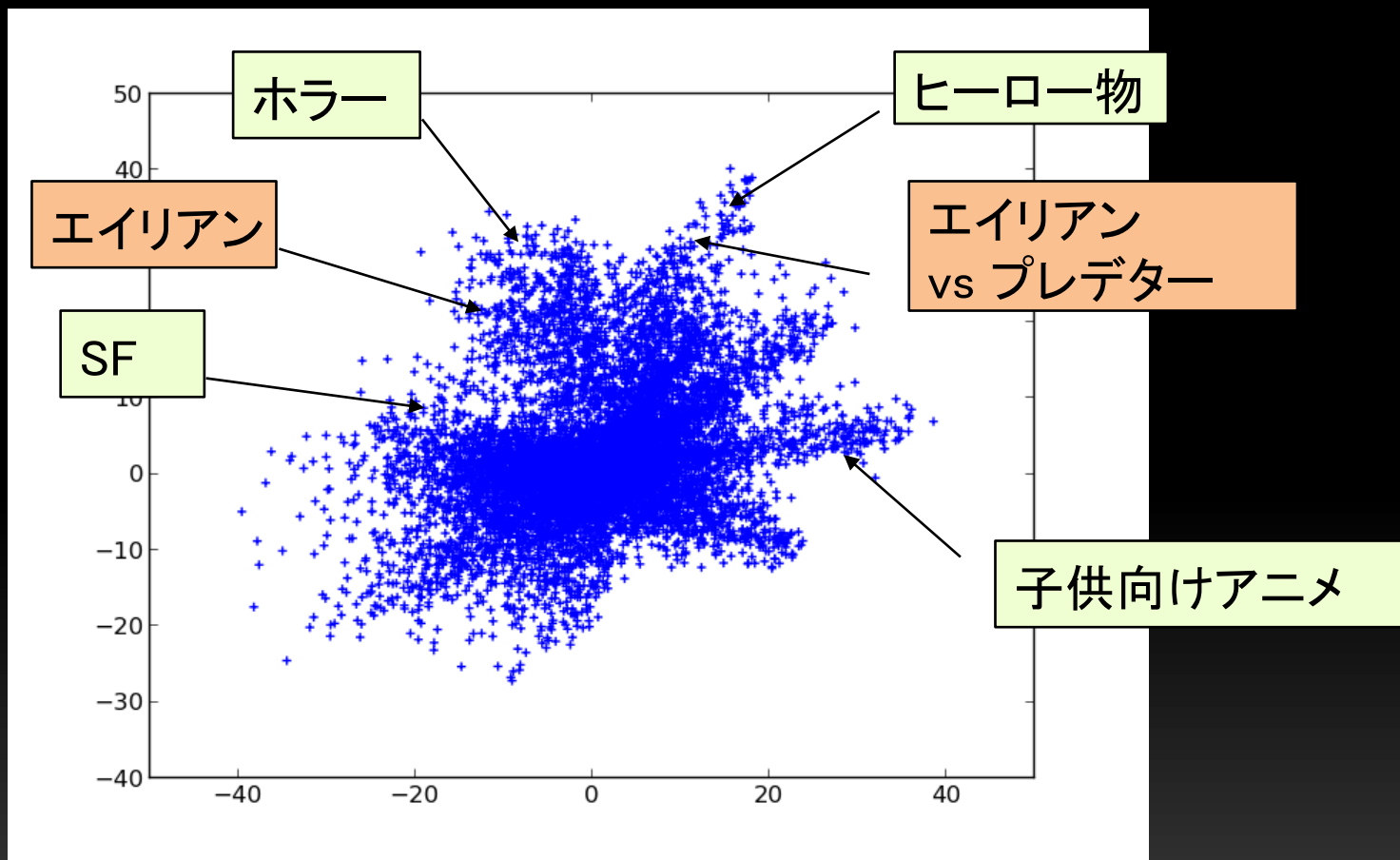
- 特徴量を自動抽出
 - 専門知識が不要
- 学習器が判別・分離しやすい特徴量
 - 応用もしやすいのでは？

特徴空間(従来手法=PCA)



- 映画のタグ
 - タグ例：animal、fun、 G.Orwell、 etc
- 1000次元 → 2次元化

特徴空間 (Deep Learning)



- いい具合に特徴を取得
- いい具合に特徴空間に分布

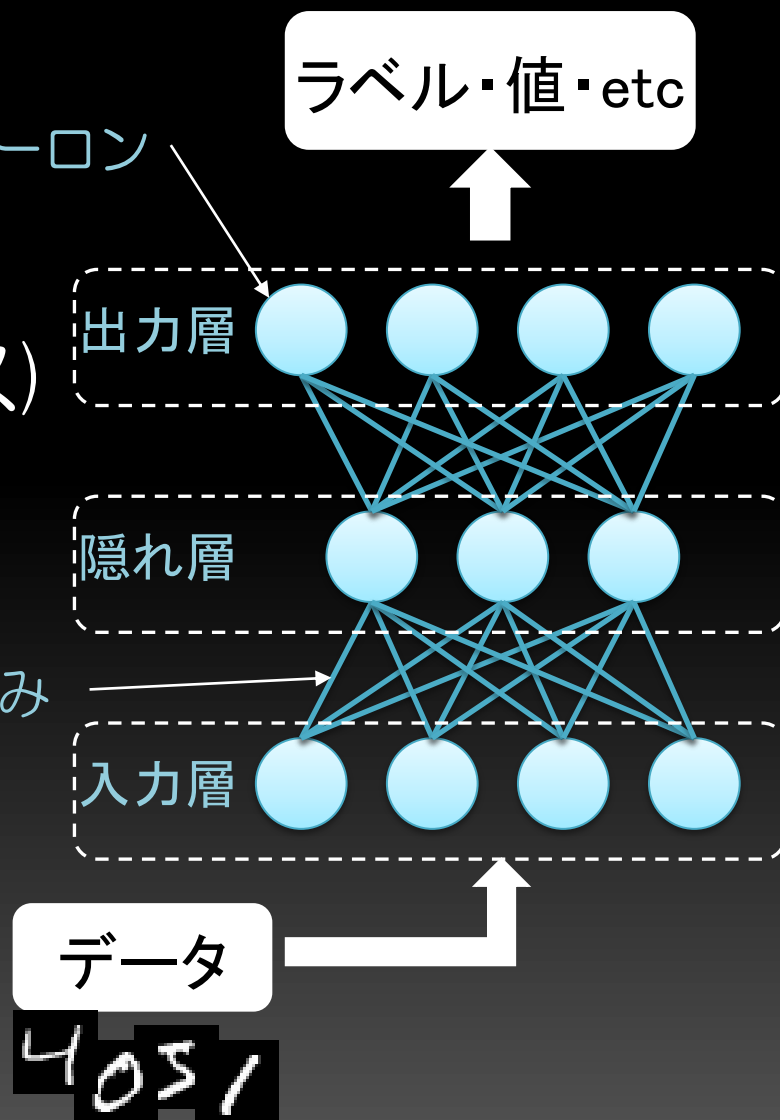
目次



- ・DeepLearningとGPU
 - ・DeepLearningとは
 - ・仕組みと計算
 - ・初期化の1手法：AutoEncoderとその応用
- ・G-DEPテストドライブ
- ・ビッグデータ・GPU・DeepLearningの接点

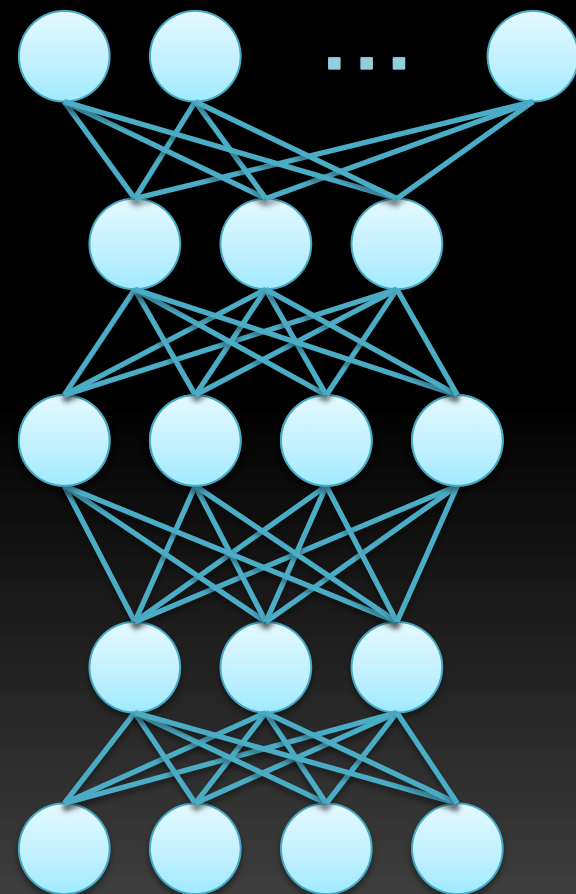
Neural Network

- 脳機能をモデル化 ニューロン
 - ニューロンの層
 - 層間で結合(シナプス)
- シナプスの結合重みを変えることにより学習



DeepLearningの仕組み

- ・DL=多層NN
- ・例：Deep CNN
 - ・8層
 - ・55万ニューロン
 - ・6000万結合
- ・ニューロン・結合が多
 - ・計算量が多
 - ・良い学習手法無し

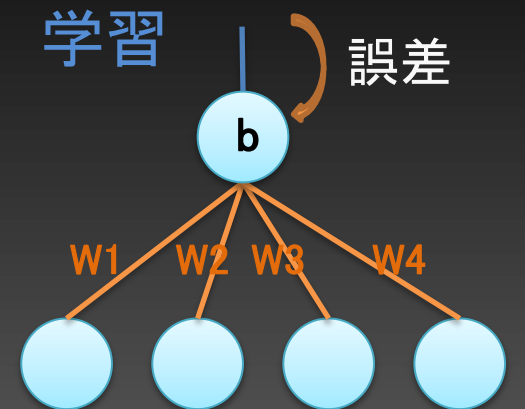
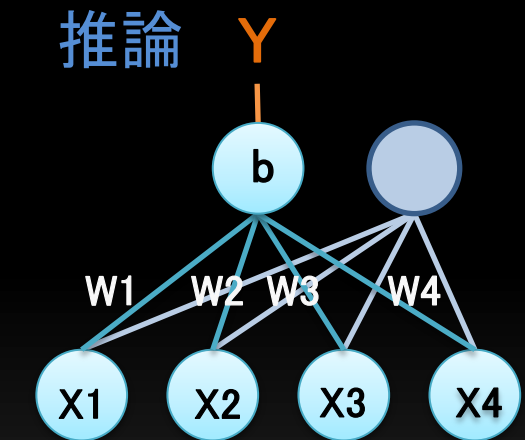


計算手法の進化：GPU

- 推論： $Y = f(\sum W_i * X_i + b)$, f : 非線形関数
- 学習：誤差逆伝播 + 重み更新
- 計算の特性：
 - 層毎の計算 → 並列可
 - 積和が中心

総積和数 = 前の層のNeuron数
 × 対象の層のNeuron数
 × 入力データ数
 × 学習回数

→ GPUによる高速化

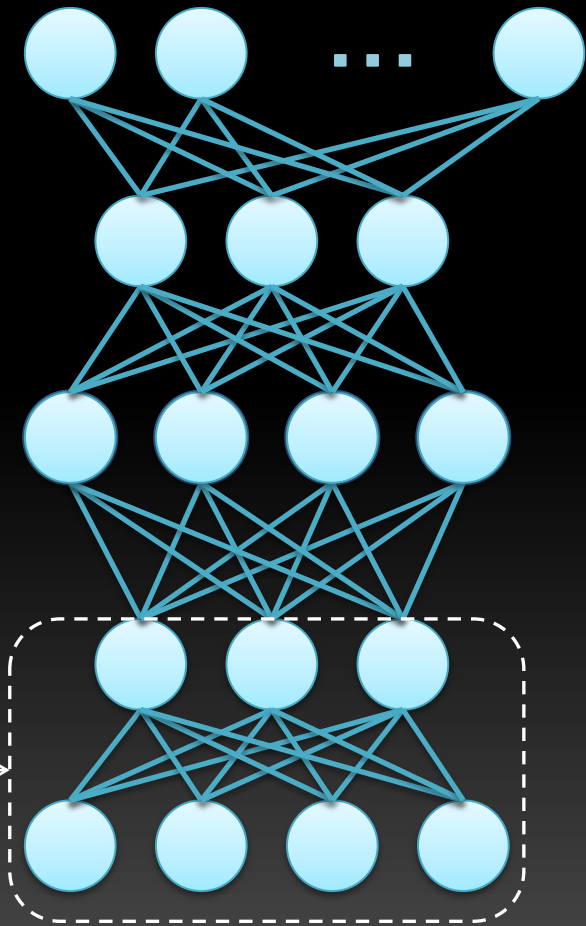
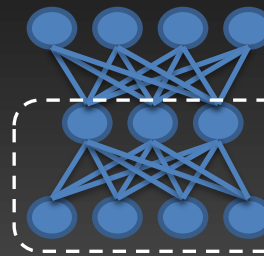


学習手法の進化

- ・ランダムな初期値で開始
→うまく学習しない
- ・「良い初期値」
 - ・層毎に重みを初期化
(pre-training)
 - ・最後に全体として学習
(fine-tuning)

AutoEncoder

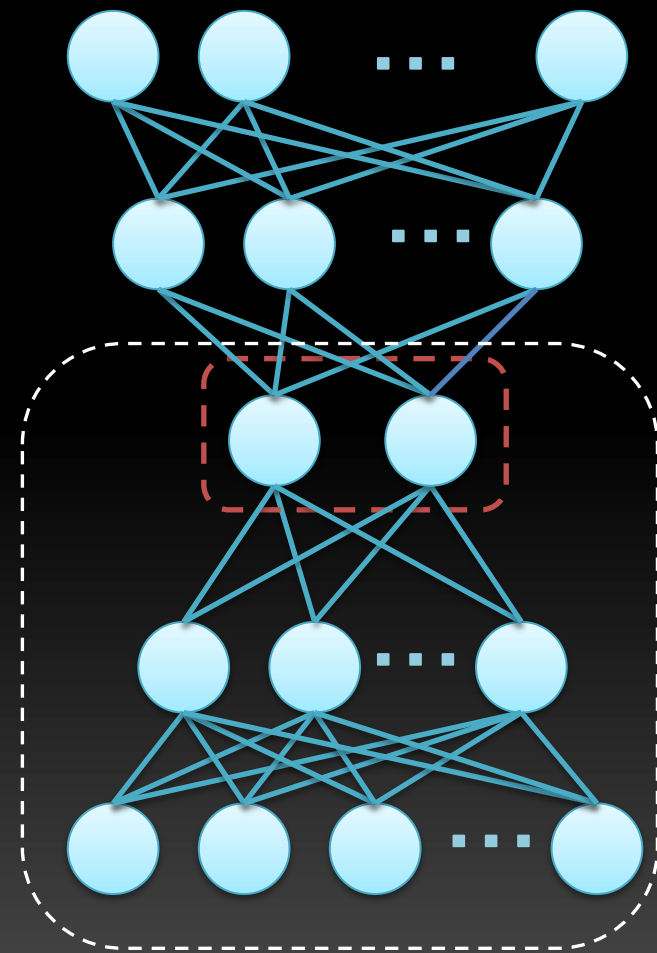
- ・上下対称NN
- ・入力=出力となるよう学習



Deep AutoEncoderと特徴空間



- 多層版AutoEncoder
 - 教師無しで学習可能
- “くびれ”
 - 入力層より小
→次元圧縮
 - 入力層より大
→Sparse Coding
→ある種の次元圧縮
- 非線形な次元圧縮器(白樫)



DeepLearningまとめ



- 多層NNによる機械学習
- 特徴量を自動抽出
- 積和計算が大量に発生 → GPU
 - 並列化可能
 - FMAにより積和は1 clock
- 良い初期値から学習を始める
- Deep AutoEncoderによる教師無し学習で次元圧縮できる

目次



- ・DeepLearningとGPU
- ・G-DEPテストドライブ
- ・ビッグデータ・GPU・DeepLearningの接点

G-DEPテストドライブ



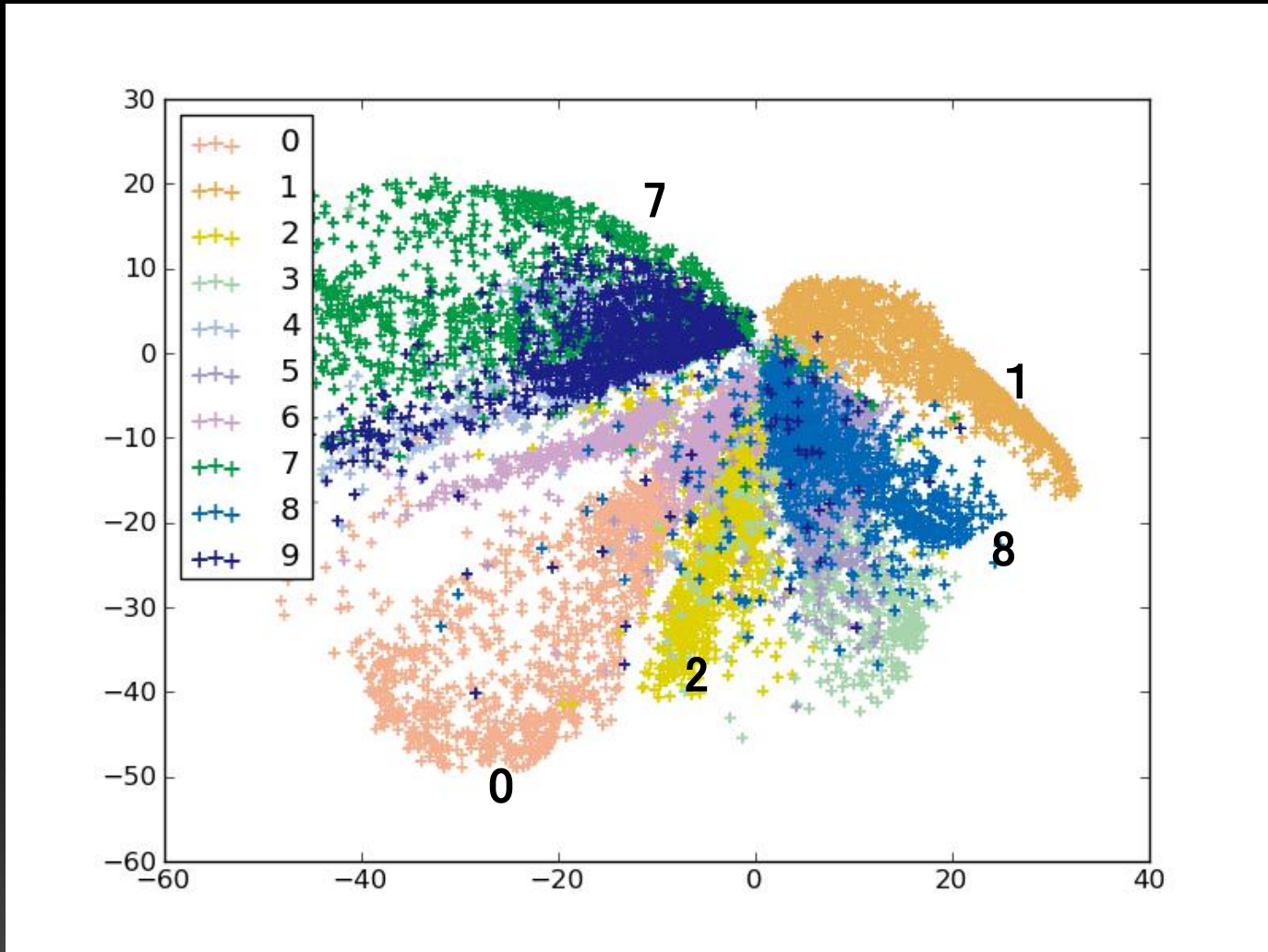
- ・Tesla K40 / K20 (マルチGPU)
 - ・Intel® Xeon E5-2650v2 2.60GHz
 - ・128GB Memory
 - ・多彩なミドルウェアが利用可能
 - ・Linux/Windows
- ・詳細はお手元の資料をご確認ください

デモ～TheanoでDAE

- ・Hinton(2006)の再現実験
- ・Deep AutoEncoderを使い次元圧縮。
784-1000-500-25-2
- ・28x28の手書き文字、5万枚
- ・学習回数
 - ・Pre-training 15回(×4層)
 - ・Fine-Tuning 5回
- ・圧縮後の2次元コードを表示
- ・CPU版との性能比較

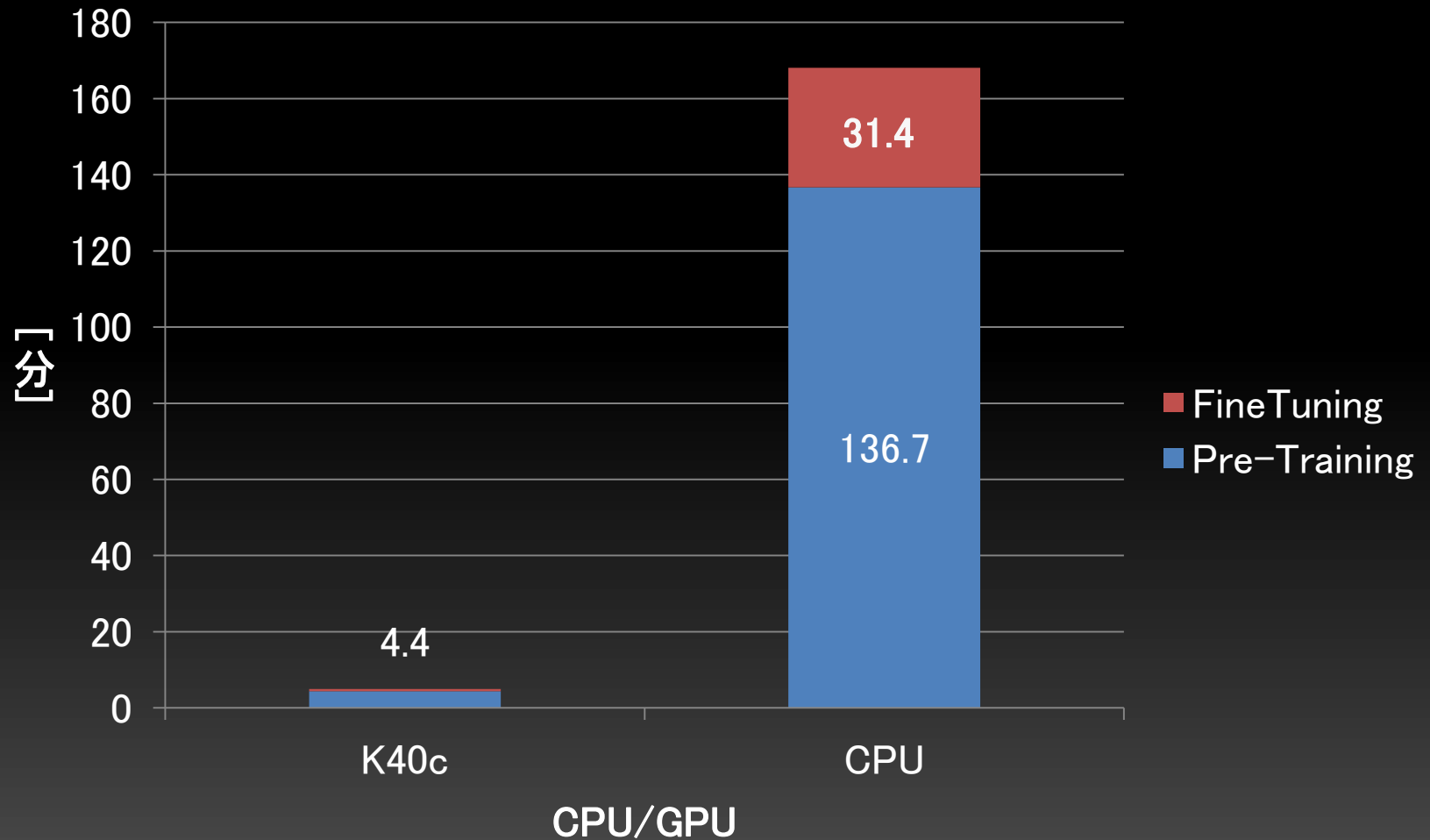
- ・Theanoとは？
 - ・PythonによるDeepLearning実装
 - ・CPU/GPUを透過的に利用可
 - ・シンボル微分に対応
 - ・pylearn2の基礎にもなっている
- ・その他実装系もあります…Let's 相談!
 - ・cuda-convnet、EBlearn、etc
 - ・http://deeplearning.net/software_links/

AutoEncoderによる圧縮コード



手書き文字(MNIST)を784-1000-500-250-2で圧縮

性能比較



目次



- ・DeepLearningとGPU
- ・G-DEPテストドライブ
- ・ビッグデータ・GPU・DeepLearningの接点

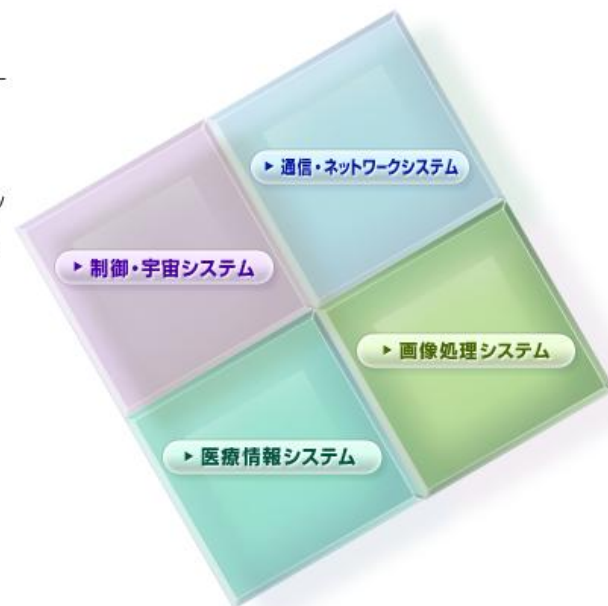
今更ですが…弊社事業分野



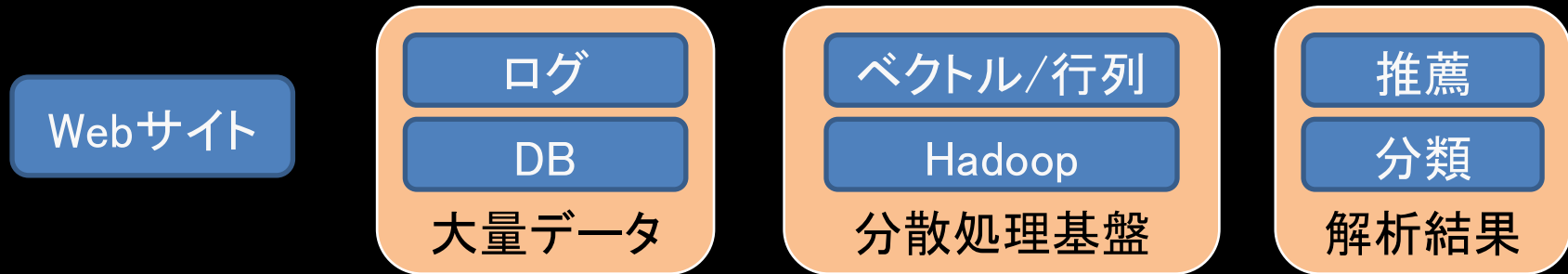
- 画像処理・通信・宇宙・医療
- 大規模メッセージ基盤
- ユーザデータ解析
- レコメンドシステム
- etc

事業内容

- ◆ソフトウェア開発
- ◆システム開発
- ◆システムインテグレーション
- ◆コンサルティング
- ◆技術開発
- ◆製品開発



BigData + GPU



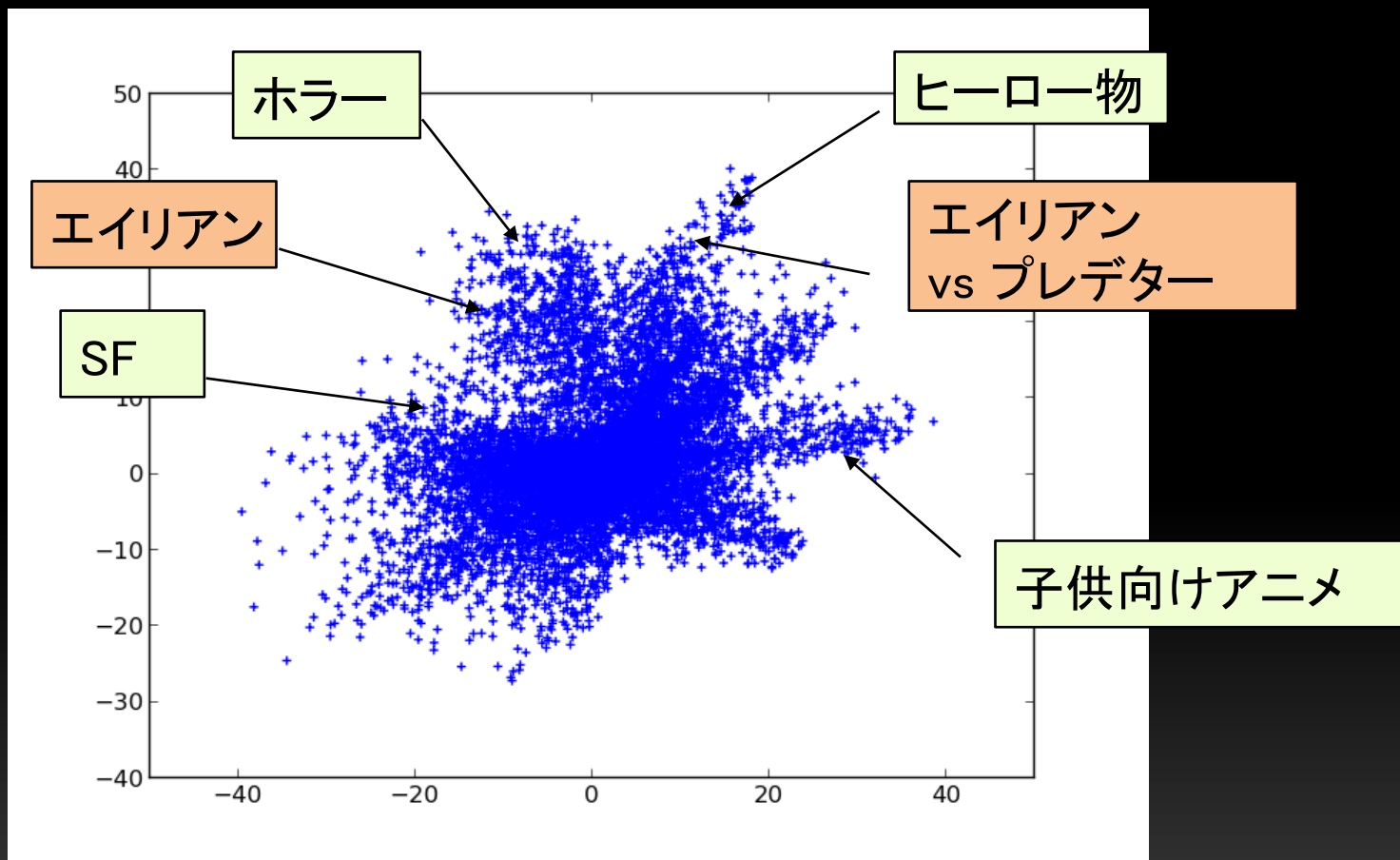
- 大量の積和を計算
- スケールアウト
 - クラスタ運用
 - MapReduceに則ったプログラム開発
- 環境面/開発面の大変さ
- JBOD(+ cache) + GPUによるスモールスタートもありでは？

BigData+DeepLearning



- 統計手法 → 機械学習。次は DeepLearningでは？
- 取り扱う対象の拡大
 - ハンドクラフトな特徴抽出が困難
 - 教師データが無い
- D.L.による特徴抽出
 - Deep AutoEncoder

[再掲]特徴空間(Deep Learning)



- 映画の口コシ情報
- 1000次元 → 2次元化

おまけ



- ブースでお待ちしております。
 - G-DEP (日本GPUコンピューティングパートナーシップ)
 - システム計画研究所 / ISP
- 技術公開サイト『**技ラボ**』でも発信していきます
 - <http://wazalabo.com/>