

CS第2 テーマ2

テーマ2の目標

データマイニングを体験する

レポート課題

課題2: 毒キノコの法則を発見せよ

内容

1. はじめに:
→ データマイニングのお話し
2. レポート課題2の説明 & 演習



1. はじめに

りやく
ご利益は？

コンピュータに載せる

スパコン



⇒ 情報が見えてくる

情報

計算

計算 → 情報を形にする

1. はじめに

計算に載せると情報が見えてくる



開発費 1000 億円～

計算で見えてくる「情報」

☆シミュレーション

★データマイニング

なぜ必要？

今回は、こちらが中心



計算で「情報」が見えてくるからじゃ

1. はじめに → データマイニング

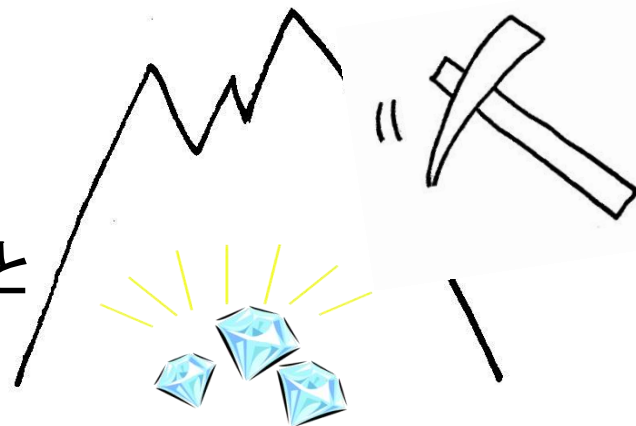
データマイニング

||

大量のデータの中から

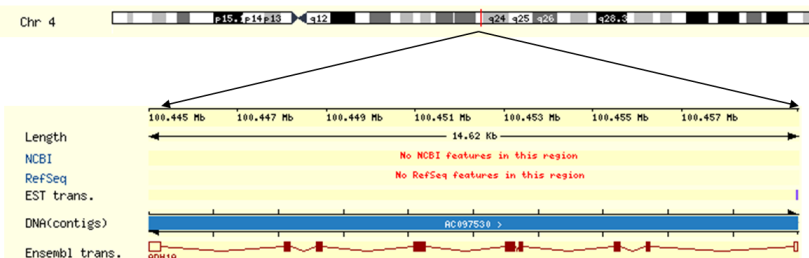
ダイヤモンド級の「情報」を発見すること

- ・ 遺伝子(ゲノム)解析
- ・ タンパク質分析 ⇒ 創薬
- ・ ビジネスデータ分析



mining = 鉱脈を発掘すること

DNA ⇒ タンパク質



↓ アミノ酸列

```
VIKCKAAVLWEVKKPFPSIEDVEVAPPKAYEVRIKMVAVGICHTDDHVVSGNLVTPLPVILGHEAAGIVES
VGEGVTTVKPGDKVIPLFTPQCGKCRVCKNPESNYCLKNDLGNPRGTLQDGRTRFTRCGKPIHHFLGTST
FSQYTVVDENAVAKIDAASPLEKVCLIGCGFSTGYGSANVAVKVPSTCAVFLGGVGLSAVMGCKAAG
AARIHAVDINKDKFAKAKELGATECINPQDYKKPIQEVLEKEMTDGGVDFSFVIGRLDTMMASLLCCHEA
CGTSVIVGVPPASQNLINPMLLLTGRTWKGAUVYGGFKSKEGIPKLVADFMKKFSLDALITHVLPFEKI
NEGFDLLHSGKSIRTVLTF
```

1. データマイニング

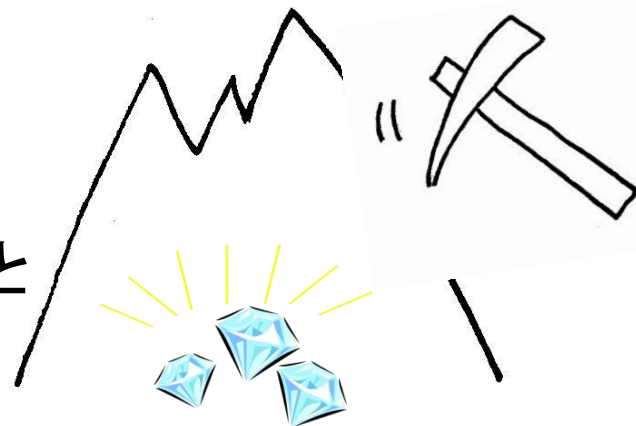
データマイニング

||

大量のデータの中から

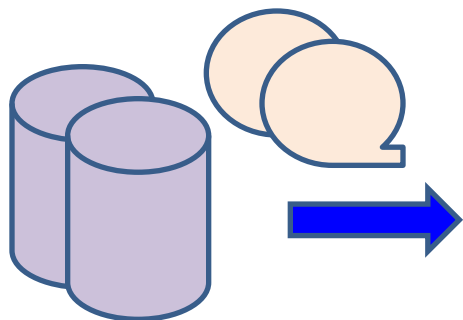
ダイヤモンド級の「情報」を発見すること

- ・ 遺伝子(ゲノム)解析
- ・ タンパク質分析 ⇒ 創薬
- ✓ ビジネスデータ分析



mining = 鉱脈を発掘すること

スーパーの購買履歴の分析



おむつ + ?

- ・ おむつ + ベビーフード
- ・ おむつ + ミルク
- ・ おむつ + **ビール !!**

1. データマイニング

法則を発見することない

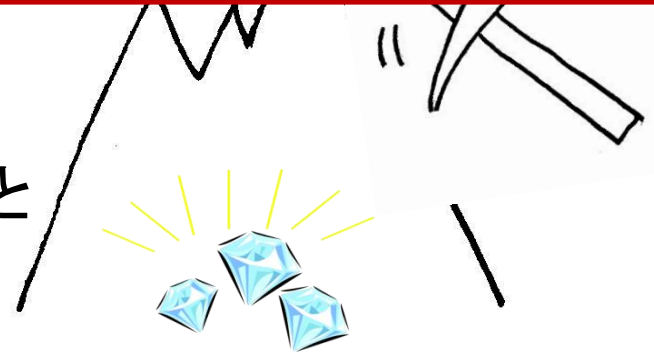
データマイニング

||

大量のデータの中から

ダイヤモンド級の「情報」を発見すること

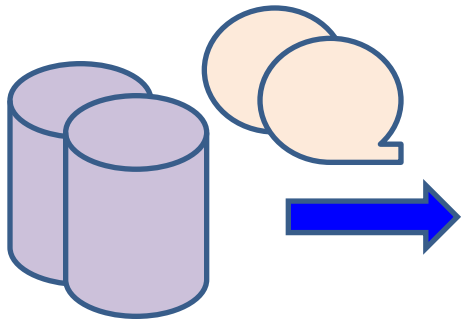
- ・ 遺伝子(ゲノム)解析
- ・ タンパク質分析 ⇒ 創薬
- ✓ ビジネスデータ分析



mining = 鉱脈を発掘すること

スーパーの購買履歴の分析

どうやって？



おむつ + ?

- ・ おむつ + ベビーフード
- ・ おむつ + ミルク
- ・ おむつ + **ビール !!**

統計？

少ないデータを生かす



大量データを生かす
計算(+統計)