# コンテンツセキュリティ実験
## - 概要説明 -

情報理工学域 Ⅱ類
渡邉洋平　阿波拓海　淺野京一

# プライバシー保護データマイニング
## （PPDM： Privacy Preserving Data Mining）

# データベースの内容：個人情報

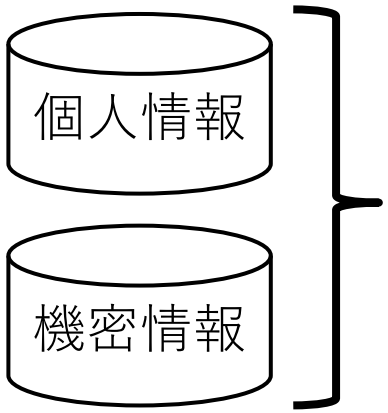
- **データベースには大量の個人情報が蓄積されている**
  - ネットショップの顧客データベース
    数100万人の顧客の氏名，年齢，性別，住所，電話番号，メールアドレス，購買履歴
  - 病院の患者データベース
    数万人の患者の氏名，年齢，性別，住所，生年月日，電話番号，病歴，医者の所見
  - 携帯電話会社の顧客データベース
    数千万人の顧客の氏名，年齢，性別，住所，電話番号，メールアドレス，携帯利用履歴，移動履歴

# データベースの内容：機密情報

- **データベースには大量の機密情報が蓄積されている**
  - 製薬会社の薬データベース
    数万種類の薬ごとに，素材，加工方法，試験結果
  - 自動車会社の生産データベース
    数百の車種毎に，部品の型式，仕入れ先，組み立て方法，強度試験
  - 計算機メーカの発明データベース
    非公開特許情報，標的にする他社の製品名，保護する自社の製品名，特許庁の審査状況

# プライバシー保護データマイニング（PPDM）

**データマイニングとは，**
**データからパターンやルールを見つけ出すこと**



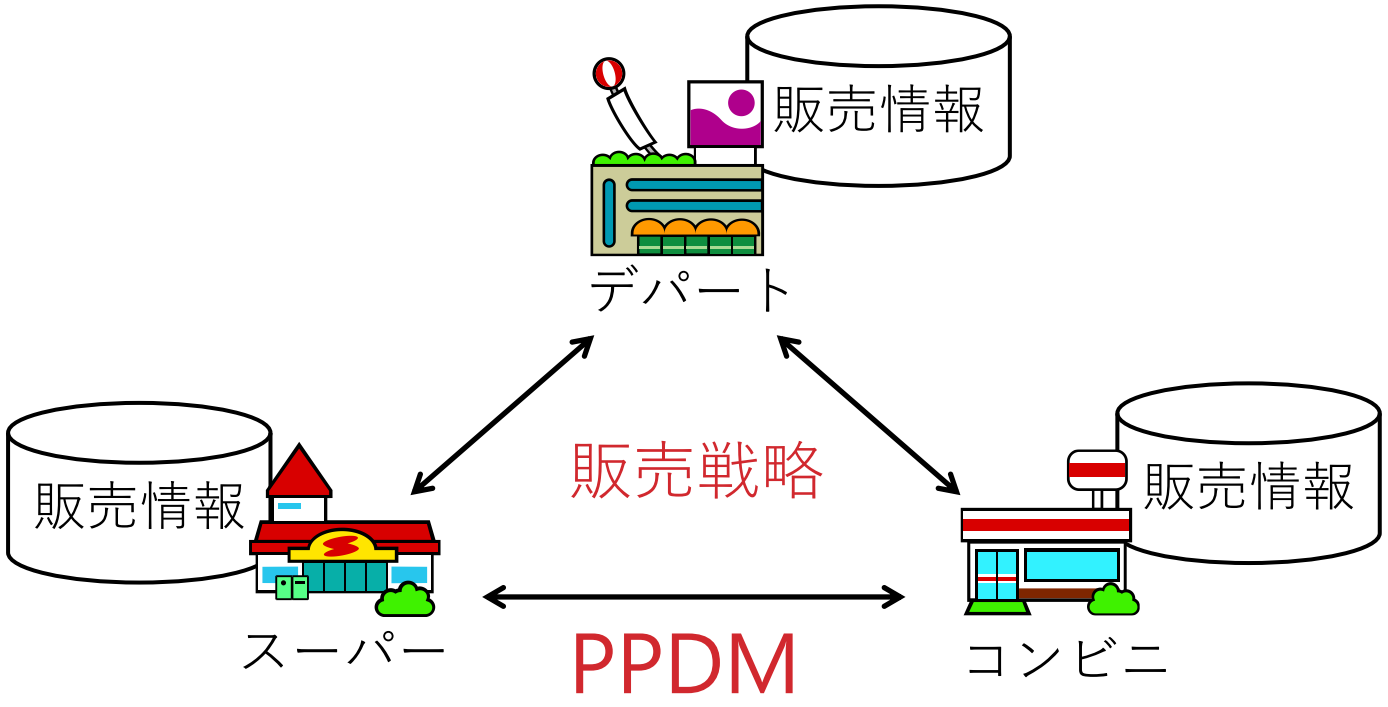

インフラ・医療・流通などの効率化
新たなサービスへつなげる

- **重要なデータは漏洩しないようプライバシー保護をする必要がある**

PPDMは，プライバシー保護をしたままデータからパターンやルールを見つけ出す技術

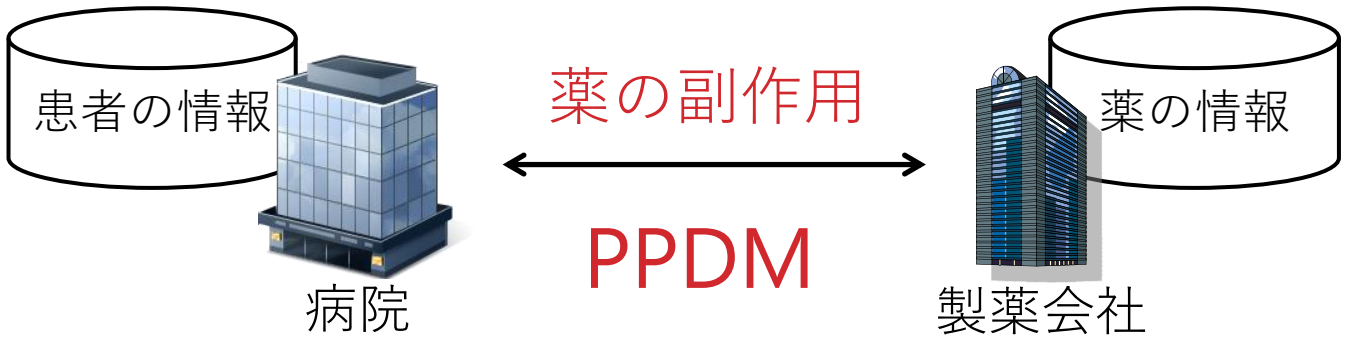
# データベースの統合利用

- **複数の組織のデータベースを統合して有用な統計情報を得たい**
- **しかし，個人情報なので法律的に組織外に出せない**
  **営業秘密なので組織外に出したくない**
- **お互いに情報を渡さないで，統合の統計情報を計算したい**



# データベースの統合利用

- **複数の組織のデータベースを統合して有用な統計情報を得たい**
- **しかし，個人情報なので法律的に組織外に出せない**
  **営業秘密なので組織外に出したくない**
- **お互いに情報を渡さないで，統合の統計情報を計算したい**



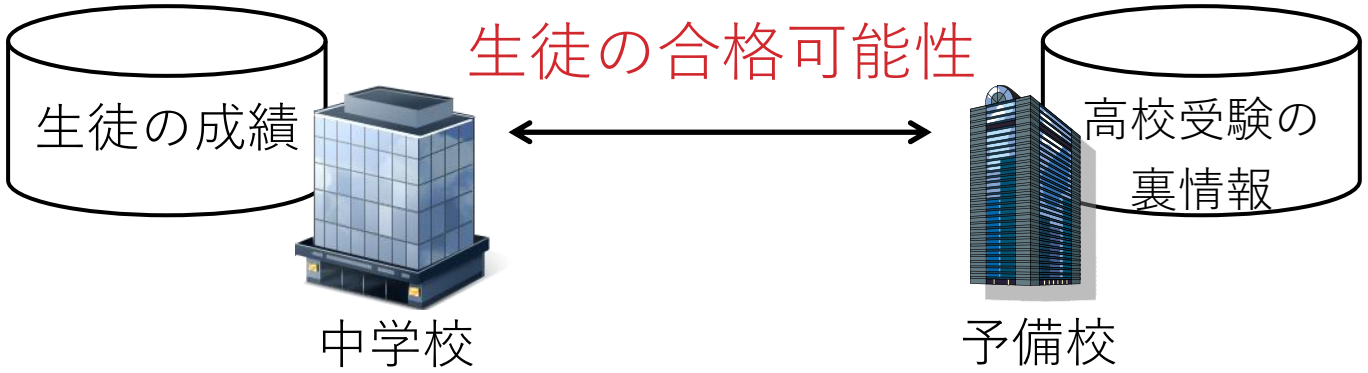

**複数の組織が自分のデータを秘密にしたままで，**
**お互いに通信し複数組織のデータを統合利用する技術**

# 課題概要

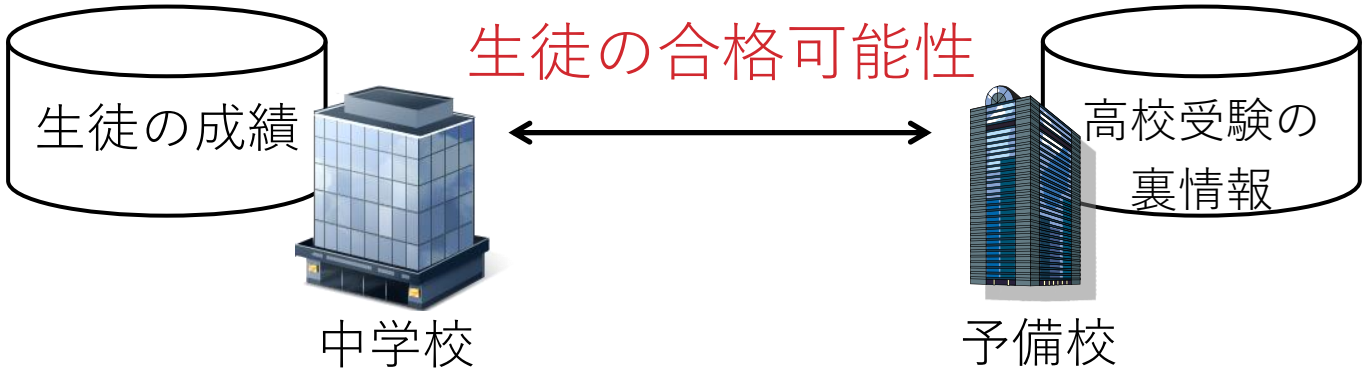
# 簡単なPPDMプログラムの作成

- **PPDM**
  **複数の組織が自分のデータを秘密にしたままでお互いに通信し，複数組織のデータを統合利用する技術である**

- **実験の課題**
  **中学校と予備校がお互いの情報を隠しながら，私立高校への合格可能性を計算する**



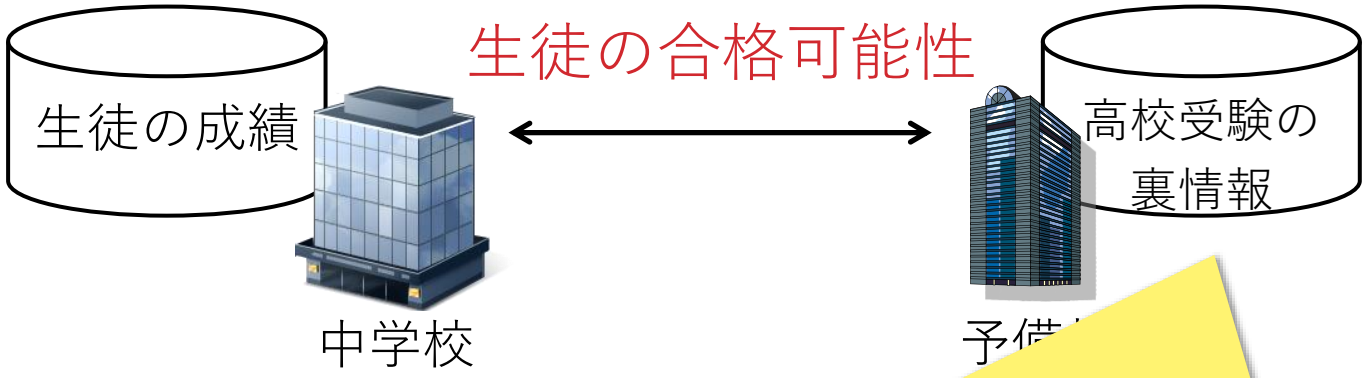
# 簡単なPPDMプログラムの作成




私立高校受験を成功させるため，
予備校のノウハウを利用したい．
しかし，生徒の成績は個人情報
なので外に出せない．

中学校に協力することで，
予備校の宣伝をしたい．
しかし，私立高校の裏情報は
貴重な営業秘密なので
外に出せない．

# 予備校の営業秘密：中学受験の裏情報

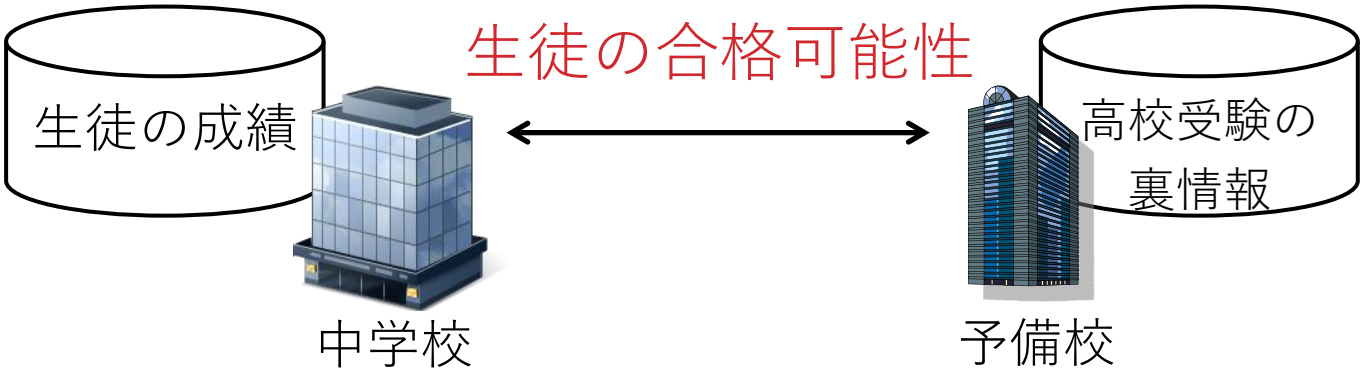



（以下はﾌｨｸｼｮﾝ）私立高校は，ボーダーラインの受験生に対して，公開された最低点とは異なる独自の判断をしている．

（例）私立A高校は，理系大学の進学実績を上げるためにボーダーラインの受験生では，数学と理科のできる生徒を合格させている場合，国語0.9倍，数学1.2倍，英語0.9倍，理科1.1倍，社会0.9倍，内申1.0の重みで判定する．

（高校受験の裏情報とは）国，数，英，理，社，内の重み

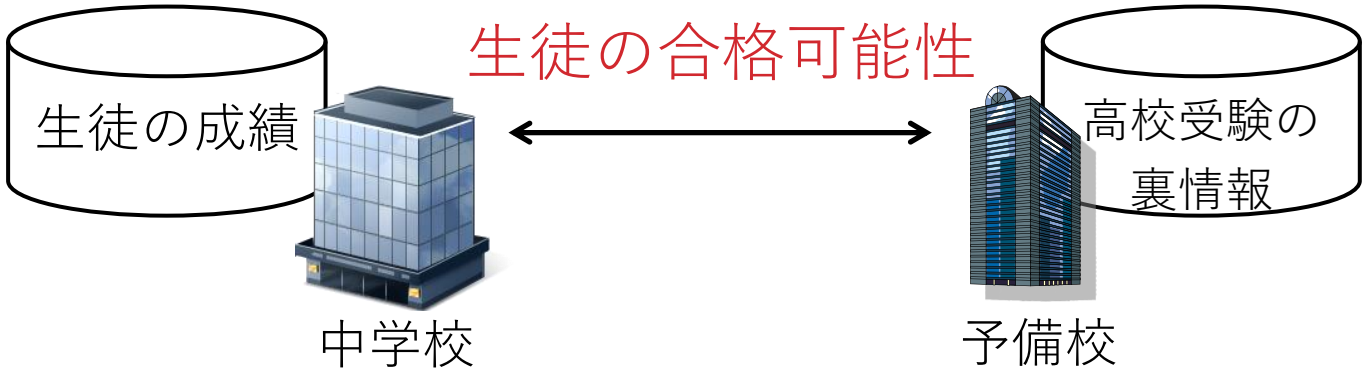
# 簡単なPPDMプログラムの作成




生徒ごとの各教科の点数

| | 国 | 数 | 英 | 理 | 社 | 内 |
|---|---|---|---|---|---|---|
| **生徒1** | 68 | 75 | 54 | 82 | 77 | 73 |
| **生徒2** | 80 | 72 | 59 | 60 | 67 | 76 |
| **生徒3** | 85 | 91 | 72 | 68 | 67 | 88 |
| **生徒4** | 71 | 65 | 95 | 75 | 60 | 80 |

# 簡単なPPDMプログラムの作成




生徒ごとの各教科の点数

| | 国 | 数 | 英 | 理 | 社 | 内 |
|---|---|---|---|---|---|---|
| **生徒1** | 68 | 75 | 54 | 82 | 77 | 73 |
| **生徒2** | 80 | 72 | 59 | 60 | 67 | 76 |
| **生徒3** | 85 | 91 | 72 | 68 | 67 | 88 |
| **生徒4** | 71 | 65 | 95 | 75 | 60 | 80 |

高校ごとの各教科の点数

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **国** | 0.9 | 1.2 | 1.1 | 1.1 |
| **数** | 1.2 | 0.8 | 1.1 | 1.1 |
| **英** | 0.9 | 1.1 | 0.9 | 0.9 |
| **理** | 1.1 | 0.8 | 0.9 | 0.8 |
| **社** | 0.9 | 1.2 | 0.9 | 0.8 |
| **内** | 1.0 | 0.9 | 1.1 | 1.3 |

# プログラムの概要

# 中学校と予備校が持つデータ

中学校：**成績行列**（生徒数×科目数）

| | 国 | 数 | 英 | 理 | 社 | 内 |
|---|---|---|---|---|---|---|
| **生徒1** | 68 | 75 | 54 | 82 | 77 | 73 |
| **生徒2** | 80 | 72 | 59 | 60 | 67 | 76 |
| **生徒3** | 85 | 91 | 72 | 68 | 67 | 88 |
| **生徒4** | 71 | 65 | 95 | 75 | 60 | 80 |

×

予備校：**重み行列**（科目数×高校数）

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **国** | 0.9 | 1.2 | 1.1 | 1.1 |
| **数** | 1.2 | 0.8 | 1.1 | 1.1 |
| **英** | 0.9 | 1.1 | 0.9 | 0.9 |
| **理** | 1.1 | 0.8 | 0.9 | 0.8 |
| **社** | 0.9 | 1.2 | 0.9 | 0.8 |
| **内** | 1.0 | 0.9 | 1.1 | 1.3 |

# 計算したい情報

中学校：**成績行列**（生徒数×科目数）

| | 国 | 数 | 英 | 理 | 社 | 内 |
|---|---|---|---|---|---|---|
| **生徒1** | 68 | 75 | 54 | 82 | 77 | 73 |
| **生徒2** | 80 | 72 | 59 | 60 | 67 | 76 |
| **生徒3** | 85 | 91 | 72 | 68 | 67 | 88 |
| **生徒4** | 71 | 65 | 95 | 75 | 60 | 80 |

×

予備校：**重み行列**（科目数×高校数）

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **国** | 0.9 | 1.2 | 1.1 | 1.1 |
| **数** | 1.2 | 0.8 | 1.1 | 1.1 |
| **英** | 0.9 | 1.1 | 0.9 | 0.9 |
| **理** | 1.1 | 0.8 | 0.9 | 0.8 |
| **社** | 0.9 | 1.2 | 0.9 | 0.8 |
| **内** | 1.0 | 0.9 | 1.1 | 1.3 |

成績行列と重み行列の積行列：**適性行列**

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **生徒1** | ??? | | | |
| **生徒2** | | | | |
| **生徒3** | | | | |
| **生徒4** | | | | |

# 計算したい情報

中学校：**成績行列**（生徒数×科目数）

| | 国 | 数 | 英 | 理 | 社 | 内 |
|---|---|---|---|---|---|---|
| **生徒1** | 68 | 75 | 54 | 82 | 77 | 73 |
| **生徒2** | 80 | 72 | 59 | 60 | 67 | 76 |
| **生徒3** | 85 | 91 | 72 | 68 | 67 | 88 |
| **生徒4** | 71 | 65 | 95 | 75 | 60 | 80 |

×

予備校：**重み行列**（科目数×高校数）

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **国** | 0.9 | 1.2 | 1.1 | 1.1 |
| **数** | 1.2 | 0.8 | 1.1 | 1.1 |
| **英** | 0.9 | 1.1 | 0.9 | 0.9 |
| **理** | 1.1 | 0.8 | 0.9 | 0.8 |
| **社** | 0.9 | 1.2 | 0.9 | 0.8 |
| **内** | 1.0 | 0.9 | 1.1 | 1.3 |

成績行列と重み行列の積行列：**適性行列**

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **生徒1** | **432.0** | | | |
| **生徒2** | | | | |
| **生徒3** | | | | |
| **生徒4** | | | | |

（例）

生徒１の高校Aに対する適合度＝

**(68×0.9)+(75×1.2)+(54×0.9)**
**+(82×1.1)+(77×0.9)+(73×1.0)** ＝**432.0**

単純な点数の和は**429**

生徒1は理系高校向きと言える

# 計算したい情報

中学校：**成績行列**（生徒数×科目数）

| | 国 | 数 | 英 | 理 | 社 | 内 |
|---|---|---|---|---|---|---|
| **生徒1** | 68 | 75 | 54 | 82 | 77 | 73 |
| **生徒2** | 80 | 72 | 59 | 60 | 67 | 76 |
| **生徒3** | 85 | 91 | 72 | 68 | 67 | 88 |
| **生徒4** | 71 | 65 | 95 | 75 | 60 | 80 |

×

予備校：**重み行列**（科目数×高校数）

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **国** | 0.9 | 1.2 | 1.1 | 1.1 |
| **数** | 1.2 | 0.8 | 1.1 | 1.1 |
| **英** | 0.9 | 1.1 | 0.9 | 0.9 |
| **理** | 1.1 | 0.8 | 0.9 | 0.8 |
| **社** | 0.9 | 1.2 | 0.9 | 0.8 |
| **内** | 1.0 | 0.9 | 1.1 | 1.3 |

成績行列と重み行列の積行列：**適性行列**

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **生徒1** | **432.0** | | | |
| **生徒2** | | | | |
| **生徒3** | | | | |
| **生徒4** | | | | |

**合格判定**

**合格最低点**

| 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|
| 430 | 440 | 420 | 390 |

# 中学校と予備校が持つデータ

中学校：**成績行列**（生徒数×科目数）

| | 国 | 数 | 英 | 理 | 社 | 内 |
|---|---|---|---|---|---|---|
| **生徒1** | 68 | 75 | 54 | 82 | 77 | 73 |
| **生徒2** | 80 | 72 | 59 | 60 | 67 | 76 |
| **生徒3** | 85 | 91 | 72 | 68 | 67 | 88 |
| **生徒4** | 71 | 65 | 95 | 75 | 60 | 80 |

×

予備校：**重み行列**（科目数×高校数）

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **国** | 0.9 | 1.2 | 1.1 | 1.1 |
| **数** | 1.2 | 0.8 | 1.1 | 1.1 |
| **英** | 0.9 | 1.1 | 0.9 | 0.9 |
| **理** | 1.1 | 0.8 | 0.9 | 0.8 |
| **社** | 0.9 | 1.2 | 0.9 | 0.8 |
| **内** | 1.0 | 0.9 | 1.1 | 1.3 |

合否をいれてください

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **生徒1** | 合 | ??? | | |
| **生徒2** | | | | |
| **生徒3** | | | | |
| **生徒4** | | | | |

**合格判定**

成績行列と重み行列の積行列：**適性行列**

| | 高校A | 高校B | 高校C | 高校D |
|---|---|---|---|---|
| **生徒1** | **432.0** | | | |
| **生徒2** | | | | |
| **生徒3** | | | | |
| **生徒4** | | | | |

# 習得する技術

- プライバシー保護の考え方
- プライバシー保護データマイニング（PPDM）の基礎
- セキュリティのための通信プロトコルの基礎
- ネットワークプログラミングの初歩