AI（人工知能）概論【Ⅱ】～教員のための実践的データサイエンス入門～

第 1 講：データサイエンスとは何か

講師：白水 始 （国立教育政策研究所 初等中等教育研究部・部長、教育データサイエンスセンター・副センター長）

はじめに

皆さんこんにちは、白水と申します。 「AI 概論Ⅱ：教員のための実践的データサイエンス入門」の第 1 講として、「データサイエンスとは何か」というテーマでお話しします 。

本講義では、教育現場の質向上につながるデータサイエンスの在り方について、その概要を紹介します 。データサイエンス一般論ではなく、聴衆である先生方を想定し、データサイエンスの中でも「教育」を主軸に考えます。なぜなら、教育とは一人一人の児童生徒の学びと育ちを支える人間的な営みだからです 。 そこにデータや分析・可視化がどう役立つのか、そしてこの営みにデータサイエンスを導入する際の留意点は何かを考えていきます 。

本講義の学習到達目標は以下の通りです 。

1. 一般的なデータサイエンスについて理解したうえで、教育のためのデータサイエンスの在り方を説明できる 。

2. データサイエンスを支える統計と機械学習という二本の柱について、具体例をもとに説明で

きる 。

3. 本講座全体の流れを理解し、学習の見通しと動機づけを持つことができる 。

---

1. 「教育データ」の定義

1.1 そもそもデータとは何か

まず、そもそも「データ」とは何かを考えてみましょう。それは「人が直面する課題に対して判断、議論、意思決定するために用いることができる情報」全般を指します 。

1.2 教育データ：本来は教育場面で得られるデータ全般

本来、「教育データ」とは教育場面で得られるデータ全般を意味します 。 例えば、教師が児童のプリントを集めて学習到達度を把握するために参照する「記述回答」や、気になった児童を後で確認するために見る「表情」や「つぶやき」、さらには授業改善のために研究者が記録した「ビデオデータ」なども、すべて教育データと言えます 。

1.3 文部科学省が定義する「教育データ」

一方で、昨今「教育データ」と呼ばれるものは、文部科学省の定義にあるように、主に「デジタルデータ」を意味することが多くなっています 。 これまではアナログな形でのデータ活用が行われてきましたが、GIGA スクール構想による 1 人 1 台端末の整備に伴い、端末の利用ログやデジタルドリルの回答時間など、紙媒体では得られなかった子供の学びに関するデジタルデータが利活用できるようになってきています 。

---

2. 「教育データ」の種類

教育データを分類すると、いくつかの種類に分けられます。 まず、業務の性質による分類として「行政系データ」「校務系データ」「学習系データ」があります 。本講座で主に対象とするのは「学習系データ」です。

また、形式による分類として「アナログデータ」と「デジタルデータ」があります 。アナログデータは連続的に変化するもの、デジタルデータは離散的に（飛び飛びで）変化するものを指します 。

さらに、「質的データ」と「量的データ」という区分も重要です 。

- 質的データ：文字、画像、動画情報など、そのままでは四則演算が適用しにくいデータ 。
- 量的データ：数値情報のデータであり、統計処理が可能になるもの 。

本講座で主に扱うのは、後者の「量的データ（構造化されたデータ）」です。前者の「質的データ（非構造化データ）」については、機械学習等で分析が可能になるという話を後半の講義で行います。

---

3. 「教育データ」利活用のメリット

教育データの利活用には、立場ごとに以下のようなメリットが考えられます 。

- 子供：自らの学びを振り返ったり、広げたり、伝えたりすることが可能になる 。

- 教師：よりきめ細かい指導や支援が可能となり、自身の経験や知見と照合することで成長につながる 。
- 保護者：子供の学校での様子を確認するなど、学校との連携が容易になる 。
- 学校設置者：類似の自治体との比較や施策の改善が容易になる 。

ただし重要なのは、これらすべての基盤に「学習過程の解明」があるということです 。子供や教師がどう学ぶのかというプロセスが分からなければ、データを有効に活用した支援もできません。教育データをまずは学習過程の解明に役立てることが重要です。

---

4. 「データサイエンス」とは何か

データサイエンスとは何かを整理すると、以下の 3 つの視点（perspective）が重なる領域にあると言えます 。

1. Computational（計算機科学的視点）
2. Statistical（統計的視点）
3. Human（人間的視点）

これらが統合された場所に「Data Science」が存在します 。

---

5. 「教育データサイエンス」とは何か

この枠組みを教育分野に適用したのが「教育データサイエンス」です。国立教育政策研究所のシンポジウムで Sanne Smith 氏が示した図を用いて説明します 。

- コンピューターサイエンス（Computer Science）
- 統計（Statistics）
- 教育理論と実践・専門分野（Education Theory and Practice / Domain Expertise）

これらの重なり合いを見てみましょう 。

- コンピューターサイエンス × 教育理論 ＝ ソフトウェア開発（Software Development）
- コンピューターサイエンス × 統計 ＝ 機械学習（Machine Learning）
- 教育理論 × 統計 ＝ 従来の研究（Traditional Research）

そして、これら 3 つの円がすべて重なる中心に位置するのが「教育データサイエンス（Education Data Science）」です 。

Sanne Smith 氏は、「教育データサイエンスには光（チャンス）と影（課題）がある」と指摘しています 。データから学ぶことは重要ですが、教育におけるデータサイエンスは慎重に行われる必要があります。そのため、単なる教育学者でもデータサイエンティストでもない、「教育データサイエンティスト」を育成する必要があるのです 。

---

6. 教育に統計的な視点を持ち込む

「教育データサイエンス」を構成する要素の一つ、「統計的な視点」を教育に持ち込む例として、「多層型支援」を紹介します 。

多層型支援では、子供たちを「支援が必要／不要」と最初から二分するのではなく、通常の学級において多様な子供がいることを前提とした授業・学級づくり（第 1 層支援）を行います 。その上で、第 1 層支援のみでは不十分な場合に、第 2 層、第 3 層と支援を徐々に付け足し、その有効性を評価・改善していきます 。 「できない」原因を子供個人に帰属させるのではなく、環境（指導や支援）との相互作用と捉え、教師チームで意思決定を行う点が特徴です 。

分布という視点

ある学校長は、このアプローチを導入する際、「C プラスの子供たちを B 層に確実に引き上げられる授業をしよう」と発言しました 。 これは「分布」という統計的な視点を教育現場に持ち込んだ例と言えます 。 通常、特定の軸（例えば成績）で子供たちを見ると、正規分布のような形になります（A 層、B 層、C 層） 。教員は、極端に進んでいる子や遅れている子に目が行きがちですが、ボリュームゾーンである中間層（B 層）への意識が薄くなることがあります。 「クラス全体（第 1 層)」への支援を強化し、分布全体を引き上げることで、結果的に個別の支援が必要な子供たち（第 2 層・第 3 層）への対応も明確になります 。このように「分布」の視点を持つことは、教育実践において新たな気づきを与えてくれます。

---

7. 教育に統計×機械学習の視点を持ち込む

次に、統計に加えて「機械学習」の視点を持ち込む例を考えます 。 例として挙げられるのが「教育付加価値評価システム（EVAAS)」です 。これは、学校や教師が生徒の学力向上に与えた影響（付加価値）を測定するツールです 。

EVAAS では、過去の実績データから「その生徒が将来どれくらいの成績をとるか」を予測（Predicted Performance）します 。そして、実際の成績（Current Performance）が予測を上回った場合、その差分（Value-added）を「教師の指導の効果」とみなします 。 この背景には、生徒の学力の伸びから、家庭の経済状況や学校・地域の影響などの要因を統計的に取り除き、教師個人の効果を抽出するモデルが存在します 。

---

8. 教育データサイエンスの目指すべき方向性

しかし、EVAAS のようなシステムには課題もあります。実際にアメリカでは、この指標を教員の給与や賞与に直結させた結果、訴訟や現場の混乱を招いた事例があります。教育の成果を学力テストの結果だけで判断することへの懸念や、教員間の分断を招く恐れがあるからです。

Sanne Smith 氏が述べたように、私たちは教育データサイエンスに対して慎重であるべきです 。教育データサイエンスは、教育をより良くするためのものであり、教育現場を混乱させるものであってはなりません 。

避けるべき方向性（教育をおかしくする例）

- データ（結果）を基にインセンティブを付けようとする（ボトムアップな管理)。

- 教育の成果を標準的な学力だけで見る（資質・能力の一体性を見落とす)。
- 教師を相対評価し、分断する。

目指すべき方向性

- 仮説（学習理論）をもとに、まずは「よい授業」を創る（トップダウンなアプローチ)。
- 資質・能力を一体的に育成する視点を持つ。
- 教師の「協働」を何よりも大切にする。
- その目的のために、テクノロジーとデータを活用する。

このようなスタンスこそが、教育データサイエンスを健全に機能させるための「門番」の役割を果たすと考えます 。 詳細については、国立教育政策研究所 教育データサイエンスセンターの Web サイトもぜひご覧ください 。

課題

本講義の理解を深めるため、以下の課題について考えてみてください 。

1. 教育データと教育データサイエンスとは何か、本講座の例を結び付けて説明してください 。
2. データサイエンスを教育に導入する際の留意点を述べてください 。

以上で第 1 講を終わります。

AI（人工知能）概論【Ⅱ】～教員のための実践的データサイエンス入門～

第 1 講：データサイエンスとは何か

講師：白水　始（国立教育政策研究所　初等中等教育研究部・部長、教育データサイエンスセンター・副センター長）

はじめに

皆さんこんにちは、白水と申します。「AI 概論Ⅱ：教員のための実践的データサイエンス入門」の第 1 講として、「データサイエンスとは何か」というテーマでお話しします 。

本講義では、教育現場の質向上につながるデータサイエンスの在り方について、その概要を紹介します 。データサイエンス一般論ではなく、聴衆である先生方を想定し、データサイエンスの中でも「教育」を主軸に考えます。なぜなら、教育とは一人一人の児童生徒の学びと育ちを支える人間的な営みだからです 。 そこにデータや分析・可視化がどう役立つのか、そしてこの営みにデータサイエンスを導入する際の留意点は何かを考えていきます 。

本講義の学習到達目標は以下の通りです 。

1. 一般的なデータサイエンスについて理解したうえで、教育のためのデータサイエンスの在り方を説明できる 。
2. データサイエンスを支える統計と機械学習という二本の柱について、具体例をもとに説明できる 。
3. 本講座全体の流れを理解し、学習の見通しと動機づけを持つことができる 。

1.「教育データ」の定義

1.1 そもそもデータとは何か

まず、そもそも「データ」とは何かを考えてみましょう。それは「人が直面する課題に対して判断、議論、意思決定するために用いることができる情報」全般を指します 。

1.2 教育データ：本来は教育場面で得られるデータ全般

本来、「教育データ」とは教育場面で得られるデータ全般を意味します 。 例えば、教師が児童のプリントを集めて学習到達度を把握するために参照する「記述回答」や、気になった児童を後で確認するために見る「表情」や「つぶやき」、さらには授業改善のために研究者が記録した「ビデオデータ」なども、すべて教育データと言えます 。

1.3 文部科学省が定義する「教育データ」

一方で、昨今「教育データ」と呼ばれるものは、文部科学省の定義にあるように、主に「デジタルデータ」を意味することが多くなっています 。 これまではアナログな形でのデータ活用が行われてきましたが、GIGA スクール構想による 1 人 1 台端末の整備に伴い、端末の利用ログやデジタルドリルの回答時間など、紙媒体では得られなかった子供の学びに関するデジタルデータが利活用できるようになってきています 。

2.「教育データ」の種類

教育データを分類すると、いくつかの種類に分けられます。 まず、業務の性質による分類として「行政系データ」「校務系データ」「学習系データ」があります 。本講座で主に対象とするのは「学習系データ」です。

また、形式による分類として「アナログデータ」と「デジタルデータ」があります 。アナログデータは連続的に変化するもの、デジタルデータは離散的に（飛び飛びで）変化するものを指します 。

さらに、「質的データ」と「量的データ」という区分も重要です 。

- 質的データ：文字、画像、動画情報など、そのままでは四則演算が適用しにくいデータ 。
- 量的データ：数値情報のデータであり、統計処理が可能になるもの 。

本講座で主に扱うのは、後者の「量的データ（構造化されたデータ）」です。前者の「質的データ（非構造化データ）」については、機械学習等で分析が可能になるという話を後半の講義で行います。

---

3. 「教育データ」利活用のメリット

教育データの利活用には、立場ごとに以下のようなメリットが考えられます 。

- 子供：自らの学びを振り返ったり、広げたり、伝えたりすることが可能になる 。
- 教師：よりきめ細かい指導や支援が可能となり、自身の経験や知見と照合することで成長につながる 。
- 保護者：子供の学校での様子を確認するなど、学校との連携が容易になる 。
- 学校設置者：類似の自治体との比較や施策の改善が容易になる 。

ただし重要なのは、これらすべての基盤に「学習過程の解明」があるということです 。子供や教師がどう学ぶのかというプロセスが分からなければ、データを有効に活用した支援もできません。教育データをまずは学習過程の解明に役立てることが重要です。

---

4. 「データサイエンス」とは何か

データサイエンスとは何かを整理すると、以下の 3 つの視点（perspective）が重なる領域にあると言えます 。

1. Computational（計算機科学的視点）
2. Statistical（統計的視点）
3. Human（人間的視点）

これらが統合された場所に「Data Science」が存在します 。

---

5. 「教育データサイエンス」とは何か

この枠組みを教育分野に適用したのが「教育データサイエンス」です。国立教育政策研究所のシンポジウムで Sanne Smith 氏が示した図を用いて説明します 。

- コンピューターサイエンス（Computer Science）
- 統計（Statistics）
- 教育理論と実践・専門分野（Education Theory and Practice / Domain Expertise）

これらの重なり合いを見てみましょう 。

- コンピューターサイエンス × 教育理論 ＝ ソフトウェア開発（Software Development）
- コンピューターサイエンス × 統計 ＝ 機械学習（Machine Learning）

• 教育理論 × 統計 ＝ 従来の研究（Traditional Research）

そして、これら 3 つの円がすべて重なる中心に位置するのが「教育データサイエンス（Education Data Science）」です 。

Sanne Smith 氏は、「教育データサイエンスには光（チャンス）と影（課題）がある」と指摘しています 。データから学ぶことは重要ですが、教育におけるデータサイエンスは慎重に行われる必要があります。そのため、単なる教育学者でもデータサイエンティストでもない、「教育データサイエンティスト」を育成する必要があるのです 。

---

6. 教育に統計的な視点を持ち込む

「教育データサイエンス」を構成する要素の一つ、「統計的な視点」を教育に持ち込む例として、「多層型支援」を紹介します 。

多層型支援では、子供たちを「支援が必要／不要」と最初から二分するのではなく、通常の学級において多様な子供がいることを前提とした授業・学級づくり（第 1 層支援）を行います 。その上で、第 1 層支援のみでは不十分な場合に、第 2 層、第 3 層と支援を徐々に付け足し、その有効性を評価・改善していきます 。「できない」原因を子供個人に帰属させるのではなく、環境（指導や支援）との相互作用と捉え、教師チームで意思決定を行う点が特徴です 。

分布という視点

ある学校長は、このアプローチを導入する際、「C プラスの子供たちを B 層に確実に引き上げられる授業をしよう」と発言しました 。 これは「分布」という統計的な視点を教育現場に持ち込んだ例と言えます 。 通常、特定の軸（例えば成績）で子供たちを見ると、正規分布のような形になります（A 層、B 層、C 層） 。教員は、極端に進んでいる子や遅れている子に目が行きがちですが、ボリュームゾーンである中間層（B 層）への意識が薄くなることがあります。「クラス全体（第 1 層）」への支援を強化し、分布全体を引き上げることで、結果的に個別の支援が必要な子供たち（第 2 層・第 3 層）への対応も明確になります 。このように「分布」の視点を持つことは、教育実践において新たな気づきを与えてくれます。

---

7. 教育に統計×機械学習の視点を持ち込む

次に、統計に加えて「機械学習」の視点を持ち込む例を考えます 。 例として挙げられるのが「教育付加価値評価システム（EVAAS）」です 。これは、学校や教師が生徒の学力向上に与えた影響（付加価値）を測定するツールです 。

EVAAS では、過去の実績データから「その生徒が将来どれくらいの成績をとるか」を予測（Predicted Performance）します 。そして、実際の成績（Current Performance）が予測を上回った場合、その差分（Value-added）を「教師の指導の効果」とみなします 。この背景には、生徒の学力の伸びから、家庭の経済状況や学校・地域の影響などの要因を統計的に取り除き、教師個人の効果を抽出するモデルが存在します 。

---

8. 教育データサイエンスの目指すべき方向性

しかし、EVAAS のようなシステムには課題もあります。実際にアメリカでは、この指標を教員の給与や賞与に直結させた結果、訴訟や現場の混乱を招いた事例があります。教育の成果を学力テストの結果だけで判断することへの懸念や、教員間の分断を招く恐れがあるからです。

Sanne Smith 氏が述べたように、私たちは教育データサイエンスに対して慎重であるべきです 。教育データサイエンスは、教育をより良くするためのものであり、教育現場を混乱させるものであってはなりません 。

避けるべき方向性（教育をおかしくする例）

- データ（結果）を基にインセンティブを付けようとする（ボトムアップな管理）。
- 教育の成果を標準的な学力だけで見る（資質・能力の一体性を見落とす）。
- 教師を相対評価し、分断する。

目指すべき方向性

- 仮説（学習理論）をもとに、まずは「よい授業」を創る（トップダウンなアプローチ）。
- 資質・能力を一体的に育成する視点を持つ。
- 教師の「協働」を何よりも大切にする。
- その目的のために、テクノロジーとデータを活用する。

このようなスタンスこそが、教育データサイエンスを健全に機能させるための「門番」の役割を果たすと考えます 。 詳細については、国立教育政策研究所 教育データサイエンスセンターの Web サイトもぜひご覧ください 。

課題

本講義の理解を深めるため、以下の課題について考えてみてください 。

1. 教育データと教育データサイエンスとは何か、本講座の例を結び付けて説明してください 。
2. データサイエンスを教育に導入する際の留意点を述べてください 。

以上で第 1 講を終わります。

ご提示いただいた講義資料「AI（人工知能）概論【Ⅱ】～教員のための実践的データサイエンス入門～ 第 1 講：データサイエンスとは何か」の内容を、ご要望に基づき 5,000 字程度を目安に項目立てしてまとめます。

□

AI（人工知能）概論【Ⅱ】第 1 講：データサイエンスとは何か

Ⅰ．はじめに：教育現場のためのデータサイエンス入門

1.1 講義の目的と概要

本講義は、「AI 概論Ⅱ：教員のための実践的データサイエンス入門」の第 1 講であり、「データサイエンスとは何か」をテーマとしています。一般的なデータサイエンス論ではなく、特に**教育現場の質向上**につながるデータサイエンスの在り方に焦点を当てます。

教育とは、一人ひとりの児童生徒の「学び」と「育ち」を支える**人間的な営み**です。この営みの中で、データ、分析、可視化がどのように役立つのか、そしてデータサイエンスを導入する際の**留意点**は何かを考察します。

1.2 学習到達目標

本講義の学習到達目標は以下の通りです。

- **教育のためのデータサイエンスの理解**：一般的なデータサイエンスを理解した上で、教育のためのデータサイエンスの在り方を説明できるようになる。
- **二本の柱の理解**：データサイエンスを支える「統計」と「機械学習」という二本の柱について、具体例をもとに説明できるようになる。
- **講座全体の理解**：本講座全体の流れを理解し、今後の学習の見通しと動機づけを持つことができる。

Ⅱ．「教育データ」の定義と種類

2.1 そもそもデータとは何か

データとは、「人が直面する課題に対して**判断、議論、意思決定**するために用いることができる情報」全般を指します。

2.2 「教育データ」の定義

1. 本来の定義（教育場面で得られるデータ全般）

本来の「教育データ」は、教育場面で得られるデータすべてを意味します。これには、以下のものなどが含まれます。

- **記述回答**：教師が児童のプリントを集めて学習到達度を把握するために参照する情報。
- **観察情報**：教師が気になった児童を後で確認するために見る「表情」や「つぶやき」。
- **研究データ**：授業改善のために研究者が記録した「ビデオデータ」など。

2. 文部科学省が定義する「教育データ」（デジタルデータ）

昨今「教育データ」と呼ばれるものは、文部科学省の定義にあるように、主に**デジタルデータ**を意味することが多くなっています。

- **背景**：GIGA スクール構想による 1 人 1 台端末の整備。

- **内容**: 端末の利用ログ、デジタルドリルの回答時間など、紙媒体（アナログ）では得られなかった子供の学びに関するデジタルデータが利活用できるようになってきています。

2.3 「教育データ」の主な種類

教育データは複数の視点から分類されます。

1. 業務の性質による分類

- **行政系データ**: 文部科学省や教育委員会など、行政組織に関わるデータ。
- **校務系データ**: 学校の運営（出欠席、成績処理など）に関わるデータ。
- **学習系データ**: 児童生徒の学びの過程や成果に関わるデータ。本講座で主に扱う対象です。

2. 形式による分類

- **アナログデータ**: 連続的に変化するデータ（例：波形、アナログ時計の動き）。
- **デジタルデータ**: 離散的に（飛び飛びで）変化するデータ（例：コンピュータで処理されるデータ）。

3. 構造による分類

- **質的データ（非構造化データ）**: 文字、画像、動画情報など、そのままでは四則演算が適用しにくいデータ。
- *処理:* 後半の講義で、機械学習等により分析が可能になることが解説されます。
- **量的データ（構造化データ）**: 数値情報のデータであり、統計処理が可能になるもの。
- *処理:* 本講座で主に扱う対象です。

Ⅲ. 「教育データ」利活用のメリットと基盤

3.1 立場ごとの利活用メリット

教育データを活用することで、様々な関係者に以下のようなメリットがもたらされます。

立場	メリット
子供	自らの学びを振り返り、広げたり、伝えたりすることが可能になる。
教師	より**きめ細かい指導や支援**が可能となり、自身の経験や知見と照合することで**成長**につながる。
保護者	子供の学校での様子を確認するなど、**学校との連携**が容易になる。
学校設置者	類似の自治体との比較や**施策の改善**が容易になる。

3.2 利活用の基盤:「学習過程の解明」

上記のメリットを実現するためのすべての基盤となるのが**「学習過程の解明」**です。

- 子供や教師が**どう学ぶのか**というプロセスが分からなければ、データを有効に活用した支援はできません。
- 教育データをまずは**学習過程の解明**に役立てることが最も重要です。

Ⅳ. データサイエンスと教育データサイエンスの構造

4.1 データサイエンスの構造

一般的なデータサイエンスは、以下の 3 つの視点（perspective）が重なる領域にあると定義されます。

1. **Computational（計算機科学的視点）:** データを処理し、アルゴリズムを実行する技術的な側面。
2. **Statistical（統計的視点）:** データを分析し、傾向や予測を行う数理的な側面。
3. **Human（人間的視点）:** 課題解決や意思決定といった、データを活用する目的や文脈に関わる側面。

4.2 教育データサイエンスの定義と構造

この枠組みを教育分野に適用したものが「教育データサイエンス」です。Sanne Smith 氏のモデルに基づき、以下の 3 つの要素の重なりとして説明されます。

構成要素	領域
コンピューターサイエンス (Computer Science)	技術的な知識とスキル。
統計 (Statistics)	データ分析とモデル構築の数理。
教育理論と実践・専門分野 (Education Theory and Practice / Domain Expertise)	教育特有の知識と文脈。

4.3 各領域の重なりと「教育データサイエンス」の位置づけ

上記の 3 つの円が重なる領域は、それぞれ以下の専門分野に対応します。

重なる領域	専門分野
コンピューターサイエンス × 教育理論	ソフトウェア開発 (Software Development)
コンピューターサイエンス × 統計	機械学習 (Machine Learning)

重なる領域	専門分野
教育理論 × 統計	従来の研究 (Traditional Research)
3 つすべてが重なる中心	**教育データサイエンス** (Education Data Science)

教育データサイエンスは、単なる教育学者やデータサイエンティストでは対応できない、専門的な知識の統合が必要な分野であり、「教育データサイエンティスト」の育成が求められます。

Ⅴ. 教育における統計的視点の導入：多層型支援の例

5.1 多層型支援（RTI モデル）とは

「教育データサイエンス」を構成する要素の一つ、「統計的な視点」を教育に持ち込む具体的な例として**多層型支援**（Response to Intervention: RTI）が紹介されます。

- **前提**: 子供たちを「支援が必要／不要」と最初から二分しない。通常の学級に多様な子供がいることを前提とする。
- **実施**:
- **第 1 層支援**: 通常の学級における授業・学級づくり。
- **第 2 層、第 3 層支援**: 第 1 層支援のみでは不十分な場合に、徐々に支援を付け足す。
- **特徴**:
- 支援の有効性を**評価・改善**するプロセスを持つ。
- 「できない」原因を子供個人に帰属させず、**環境（指導や支援）との相互作用**と捉える。
- 教師チームで意思決定を行う。

5.2 「分布」という統計的な視点

ある学校長の「C プラスの子供たちを B 層に確実に引き上げられる授業をしよう」という発言は、「分布」という統計的な視点を教育現場に持ち込んだ例です。

- **正規分布の視点**: 子供たちの成績などの分布は、正規分布のような形（A 層、B 層、C 層）になることが多い。
- **現場の課題**: 教員は、極端に進んでいる子（A 層）や遅れている子（C 層）に目が行きがちで、**ボリュームゾーンである中間層（B 層）への意識**が薄くなることがある。
- **分布視点の効用**:
- 「クラス全体（第 1 層）」への支援を強化し、**分布全体を引き上げる**ことを目指す。
- 結果的に、個別の支援が必要な子供たち（第 2 層・第 3 層）への対応も明確になる。

- 「分布」の視点を持つことは、教育実践に新たな気づきを与え、**クラス全体での質の向上**に貢献します。

Ⅵ. 教育における統計 × 機械学習の導入：教育付加価値評価システム

6.1 教育付加価値評価システム（EVAAS）の例

統計に「機械学習」の視点を持ち込む例として、アメリカで導入された**教育付加価値評価システム（Education Value-added Assessment System: EVAAS）**が挙げられます。

- **目的**: 学校や教師が**生徒の学力向上に与えた影響（付加価値）**を測定するツール。
- **仕組み**:

1. 過去の実績データから、生徒が**将来どれくらいの成績をとるか**を予測（Predicted Performance）。
2. 実際の成績（Current Performance）と予測値の**差分**（Value-added）を算出。
3. この差分（付加価値）を「教師の指導の効果」とみなす。

- **背景にあるモデル**: 生徒の学力の伸びから、家庭の経済状況や学校・地域の影響などの要因を統計的に取り除くことで、**教師個人の効果**を抽出するモデルが構築されています。

Ⅶ. 教育データサイエンスの目指すべき方向性と留意点

7.1 EVAAS の課題と教訓

EVAAS のようなシステムには、教育データサイエンス導入の難しさを示す課題があります。

- **現場の混乱**: アメリカでは、この指標を教員の**給与や賞与に直結**させた結果、訴訟や現場の混乱を招きました。
- **懸念点**: 教育の成果を**学力テストの結果だけ**で判断することへの懸念や、教員間の**分断**を招く恐れがある。

Sanne Smith 氏が指摘するように、教育データサイエンスに対しては**慎重**であるべきであり、教育をより良くするためのツールとして機能させる必要があります。

7.2 避けるべき方向性（教育をおかしくする例）

データサイエンスを導入する際に、教育の本質を見失い、避けるべき方向性は以下の通りです。

- **インセンティブによる管理**: データ（結果）を基にインセンティブを付けようとする（ボトムアップな管理）。
- **成果の限定**: 教育の成果を**標準的な学力だけ**で見る（資質・能力の一体性を見落とす）。
- **教師間の分断**: 教師を相対評価し、競争を促し、分断する。

7.3 目指すべき方向性（健全な機能）

教育データサイエンスを健全に機能させるための、目指すべき方向性、つまり**「門番」**の役割を果たすスタンスは以下の通りです。

- **理論に基づく実践**: **仮説（学習理論）**をもとに、まずは**「よい授業」**を創る（トップダウンなアプローチ）。
- **一体的な育成**: 資質・能力を**一体的に育成**する視点を持つ。
- **協働の重視**: 教師の**「協働」**を何よりも大切にする。
- **目的のための活用**: 上記の目的のために、**テクノロジーとデータ**を活用する。

Ⅷ. 課題（理解度確認）

本講義の理解を深めるため、以下の課題について考察することが推奨されます。

1. **教育データと教育データサイエンスの説明**: 本講座の例を結び付けて、教育データと教育データサイエンスとは何かを説明する。
2. **データサイエンス導入の留意点**: データサイエンスを教育に導入する際の留意点を述べる。

（まとめの文字数について） 本まとめは、元の講義内容を網羅し、構造化して記述しておりますが、元のテキストの分量（約 2,500 字）から、5,000 字程度の文字数に拡大して作成することは、**内容を水増しすることになり、正確性や簡潔さを損なう**恐れがあります。

元の内容の密度を保ちつつ、講義の要点を項目立てて整理した結果、上記の構成となりました。これにより、元の講義の意図を最も明確かつ実践的に理解できる形となっています。