

金融庁 FinTech実証実験ハブ 最終報告書

2018.8.1

(申請者) 株式会社 FRONTEO
(参加金融機関) 株式会社 三菱UFJ銀行
株式会社 りそな銀行
株式会社 横浜銀行
SMBC日興証券 株式会社



1. 本実証実験の概要

1.1 本実証実験の概要および体制図

1.2 実証実験ハブ参加検討に至った背景

1.3 実証実験における目的や目指すところ

2. 実験テーマと実施方法

3. 実験条件

3.1 実験条件詳細

3.2 KIBITによる事前精度検証

4. 実験結果

4.1 結果①応接記録

4.2 結果②お客様との通話記録

5. 結論と考察

5.1 結論と考察

5.2 所感

5.3 監督指針上、確認させていただきたいポイント

5.4 実験参加者のアンケート回答内容

1.1 本実証実験の概要および体制図

KIBIT	株式会社FRONTEOが開発した、教師あり学習タイプの人工知能の名称
教師データ	人工知能に学習させるためのデータ。『正解データ』と『不正解データ』で構成
データセット	教師データに対応した、人工知能の精度評価に利用するデータ 『正解データ』と『不正解データ』で構成
学習済みモデル	『教師データ』から人工知能が学習した特定のパラメータ
観点	人工知能に学習、検知させるデータの記載内容や文脈
評価データ	評価対象のデータ群
正解データ	見つけたい正解データ群
不正解データ	見つける必要のない不正解データ群
現行方法	人が中心となって実施する確認方法
新方法【AI】	人工知能KIBITを用いて実施する確認方法

金融庁様のご指導・ご支援のもと、銀行様・証券様と協力して本実証実験を推進した。

金融庁様

- 「法令上問題がないか」「監督指針上 問題がないか」の確認、指導

FRONTEO

実証実験参加金融機関様

銀行様（3行）

三菱UFJ銀行様

りそな銀行様

横浜銀行様

証券様（1社）

SMBC日興証券様

- 実証実験実施のサポート
- データ解析、報告書作成
- 事務局業務全般

- データ準備
- 比較実験の実施
- 業務プロセスの改善検討

現行の方法とAIを用いた新方法の両方法にて、業務生産性に関する比較試験を実施した。
銀行様では「営業応接記録のチェック業務」、証券様では「通話録音記録のチェック業務」を対象業務とし、以下の検証観点にて実験を実施した。

	銀行様	証券様
実験テーマ	現行方法と新方法【AI】を用いた、 業務生産性の比較試験	
対象業務	投資信託などの金融商品販売時の 営業応接記録のチェック業務	顧客との通話時に残されるテキスト変換された 通話録音記録のチェック業務
検証観点	① 保険乗換勧誘時に、解約に伴う不利益事項に関する営業員の説明及びお客様の理解が不十分であると読み取れる記載 ② 営業員からお客様への説明で、元本割れリスクに関する説明・認識が不十分であると読み取れる記載 ③ お客様が過剰な分配金志向を有していると読み取れる記載	① お客さまから営業員へ向けられた苦情・不満の声データ

本実証実験で使用する人工知能「KIBIT」について

KIBITは『単語』ではなく『文脈』を利用して、見つけたいデータを検知します。
『文脈』を利用する事で、特定のキーワードが含まれていない文章でも検知することが可能になります。



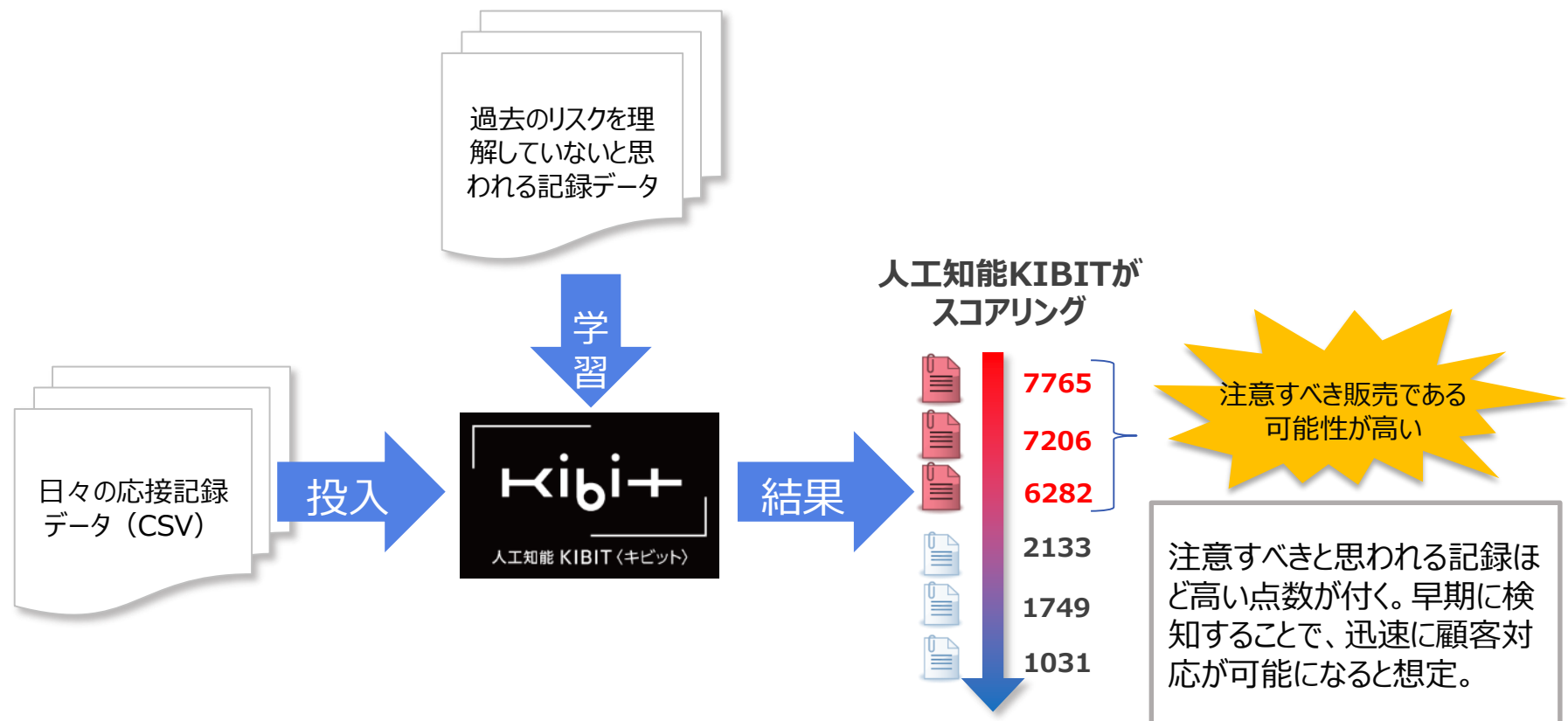
		KIBIT	一般的なAIやテキストマイニング
特長	解析対象	テキストデータ (文章データ)	AI：画像、音声、数値、テキストなど様々 テキストマイニング：単語
	学習に必要なデータ量	少ない (正解だけでなく不正解データも学習させるため数十件～数百件のデータでも学習可能)	多い (AI：正解のみ数千～数万件のデータが必要 テキストマイニング：大量の単語登録が必要)
	ユーザメリット/ デメリット	ユーザの負担が少ない (データ準備、データ処理時間、ハードウェアコストなど)	ユーザの負担が大きい (大量データ収集、長時間のデータ処理、大規模サーバコストなど)

人工知能KIBITを用いた学習モデルの作成と解析イメージ

人工知能KIBITが、事前に「見つけたいデータ」を学習しておくことで、日々の営業応接記録などのデータから見つけたい記録を効率良く、高い精度で検知できると想定している。

人工知能KIBITによるスコアリングイメージ

(例) お客様が投資のリスクを理解していない取り引きを検知させる場合



1.2 実証実験ハブ参加検討に至った背景

金融機関が金融商品をお客様に提案する際、勧誘や契約において適切な対応が求められる。現状では、営業店の役職者が注意深く応接記録内容を確認し、必要に応じて指導をしている。確認対象件数が多い場合、過去の経緯・資産状況・年齢など含めた判断は負荷が大きいことや、人による判断のブレが生じることなど、品質の更なる向上には課題がある。こういった課題への対応について、金融庁様にご指導頂きながら本実証実験を検証したい。

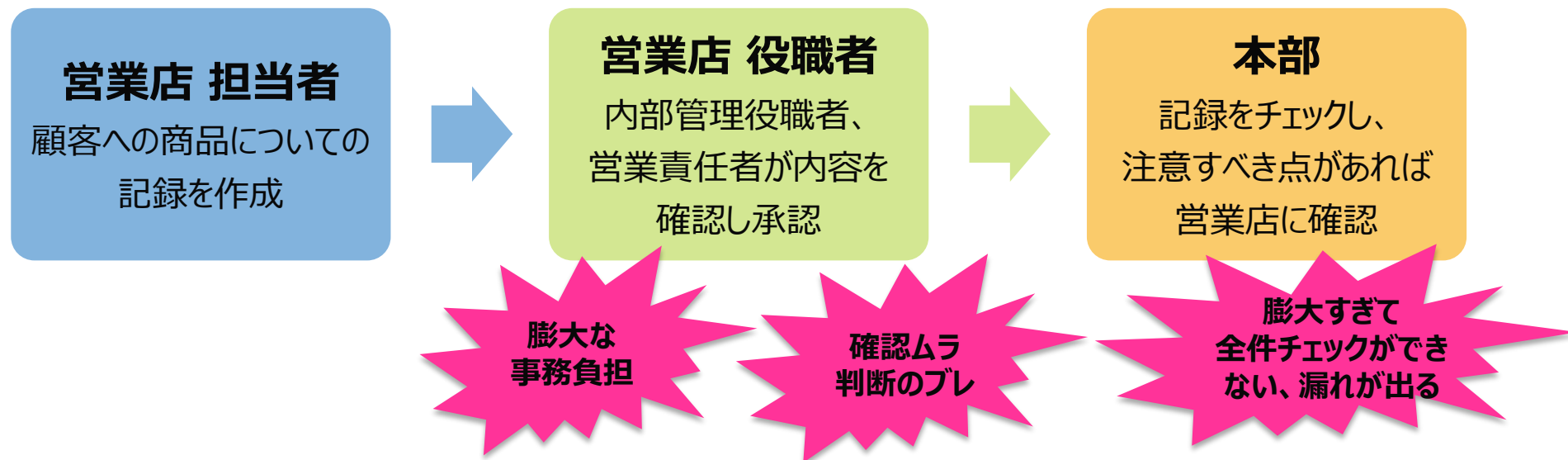
金融商品取引法の規制	
適合性の原則	<u>顧客の知識、経験、財産の状況の適切な把握が必要</u>
書面交付義務	チェックリストで未然防止可能
禁止行為	応接記録に記載なし
損失補てんの禁止	応接記録に記載なし

日本証券業協会 ガイドライン	
高齢顧客への勧誘による販売	<u>適切な説明方法や受注方法が必要</u>

応接記録などでの金融商品販売時の営業応接記録のチェック業務には、以下の大きな業務課題がある。

- ① 営業店の役職者の確認負担が大きい
- ② 役職者の知識、経験の差により、チェック品質がばらつく
- ③ 件数が多く、本部、本店での全件チェックは困難
- ④ 見落としはお客様へのリスクであり、金融機関のリスクにもなる

営業応接記録のチェック業務の流れ（イメージ）



応接記録データイメージ

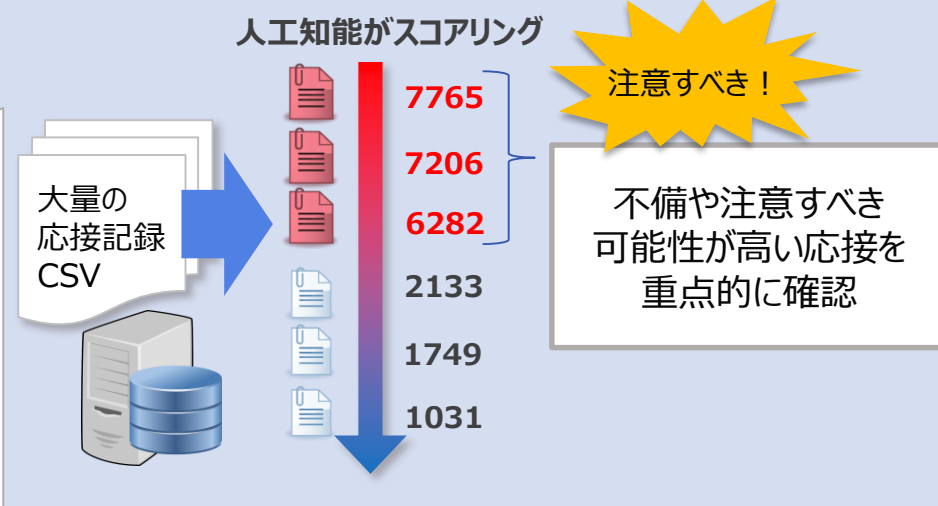
顧客A (80歳：総資産の投資比率80%)
投信B 10M解約
投信C 10M購入

【面談経緯】

現在保有銘柄の運用状況について、報告に伺ったところ「Bは益が出ているね。いつまでも持っていてつまらないから、何か別のものをやろうかな」とのご発言あったため、投信Cを案内・・・「新しいものが好きだから、この投信にしようかな」とのこと・・・Bを全額解約し、Cを購入。・・・高齢だが「頭の体操になるし、家族も承知しているよ」とのこと。投資歴も長いため・・・

人工知能KIBITによるスコアリングイメージ

人工知能が全応接記録をチェックし、不備や注意すべき順にスコアリングを行う



お客様の投資する理由が希薄であり、
行員は販売時に注意喚起すべき

以下、金融庁様HP「金融商品取引業者等向けの総合的な監督指針」より抜粋。

Ⅲ－２－５ 苦情等への対処（金融ADR制度への対応も含む。）

（１）相談・苦情・紛争等（苦情等）対処の必要性

金融商品・サービスは、リスクを内在することが多く、その専門性・不可視性等ともあいまってトラブルが生じる可能性が高いと考えられる。このため、金融商品・サービスの販売・提供に関しては、トラブルを未然に防止し顧客保護を図る観点から情報提供等の事前の措置を十分に講じることに加え、**苦情等への事後的な対処が重要となる。**

近年、金融商品・サービスの多様化・複雑化によりトラブルの可能性も高まっており、顧客保護を図り、顧客からの信頼性を確保する観点から、苦情等への事後的な対処がさらに重要になってきている。

・
・
・

Ⅲ－２－５－１ 苦情等対処に関する内部管理態勢の確立

（２）主な着眼点

③ 苦情等対処の実施態勢

□. 顧客からの苦情等について、**関係部署が連携のうえ、速やかに処理を行う態勢を整備しているか。**特に、**苦情等対処における主管部署及び担当者が、個々の職員が抱える顧客からの苦情等の把握に努め、速やかに関係部署に報告を行う態勢を整備しているか。**

苦情対処の
重要性

苦情対処の
体制について

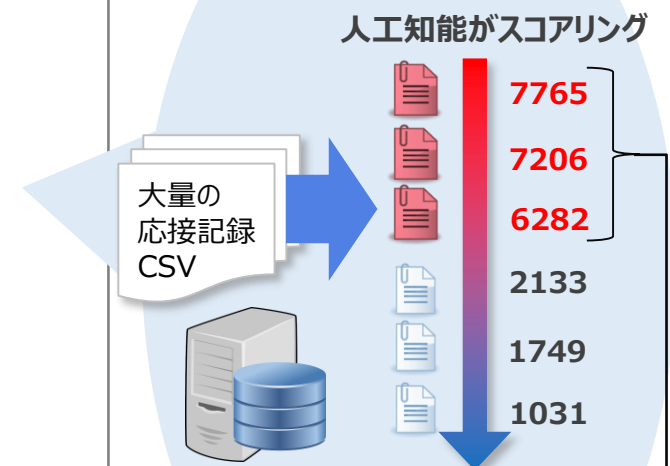
営業担当やコールセンターには日々多くのお客様からお問い合わせが寄せられる。お客様とのやり取りの通話を録音し聞き返す事で苦情などのお申し出の有無を確認することができるが、件数が多いため全てを網羅的に確認する事は現実的ではない。また、録音された通話を音声認識等で文字変換する技術が考えられるが、全てを正しく変換する事は現時点の技術では難しい。KIBITは、文字変換した中に誤変換した言葉があっても、会話の文脈全体からお客様からのお申し出や苦情に該当すると思われる会話を検知し、高いスコアを付けることで確認の優先順位付けをおこない、確認業務の高度化、効率化を実現する事ができると考えられる。

【音声通話 文字変換データイメージ】

☐ 音声認識の変換ミス

201701130930 : お客様 : A~, すみません。鈴木です。
201701130930 : 営業担当 : 鈴木様、いつもお世話になっております。
201701130930 : お客様 : この間紹介してもらった投資信託の申込み書類について聞きたいんだけど。
201701130930 : 営業担当 : はい、土曜な件でしょうか。
201701130930 : お客様 : 先日送るって言ってたけど、まだ書類が来ないよ。
音声認識の変換ミスがあっても、全体の文脈から検知
そうでしたか、お送りしたつもりだったんで図が
もしも、太平申し訳ございませんでした。
201701130930 : お客様 : これまでもこういう事があったよね？
201701130930 : 営業担当 : 本土に申し訳ありません。
201701130930 : お客様 : 何度言っても同じ事が老いるってどうなってるの？
201701130930 : 営業担当 : 念のため鈴木様の住所を確認させてもらえますか？
201701130930 : お客様 : これじゃあ心配でおたくに任せられないね。
201701130930 : 営業担当 : A~, 今後このような事の内容に注意いたします。

人工知能が全記録をチェックし、対処すべき順に苦情データをスコアリング



対処すべき可能性が高い苦情の記録を重点的に確認

1.3 本実証実験における目的や目指すところ

(人工知能KIBITを用いた実証実験を通じて整理したいと考えている論点)

1. 「金融商品販売時に営業員が記載した応接記録」や「お客さまからのお申し出や苦情」をチェック・管理する業務フローに人工知能KIBITを利用することが、金融機関の顧客管理運営上、何かしらの問題が無いかをご教示いただきたい。
2. 人工知能KIBITを活用することにより、上述の業務をより高度化できる可能性に期待している。当該業務プロセスを構築する上で、監督指針上、特に配慮すべき点があれば随時ご確認・アドバイスいただきたい。

人工知能 KIBITが金融機関のチェック業務に活用される事で、お客さまを第一に考えた業務運営の推進や長時間労働の改善、効率的・効果的な働き方の改革など、様々な社会的課題の解決につながれると期待される。

1. 顧客本位の業務運営を推進

- ・ 将来的な苦情につながりうるやり取りや、注意喚起すべき話法・勧誘手法を広範に抽出し、関係部署へ連携することにより、迅速なお客さま対応や注意喚起の高度化等を実施

2. 働き方改革の実現、人材不足への対処

- ・ 高速なチェックにより大量のモニタリング業務負荷を軽減

3. 網羅的なチェックの実現

- ・ これまでのサンプルチェックから、網羅的なチェックを実現

4. 均質なチェック態勢

- ・ 人による判断の間違い、ブレを排除

5. 顧客課題解決

- ・ チェック時間を削減する事で、お客さま対応やより付加価値のある活動に時間を活用

1. 本実証実験の概要

1.1 本実証実験の概要および体制図

1.2 実証実験ハブ参加検討に至った背景

1.3 実証実験における目的や目指すところ

2. 実験テーマと実施方法

3. 実験条件

3.1 実験条件詳細

3.2 KIBITによる事前精度検証

4. 実験結果

4.1 結果①応接記録

4.2 結果②お客様との通話記録

5. 結論と考察

5.1 結論と考察

5.2 所感

5.3 監督指針上、確認させていただきたいポイント

5.4 実験参加者のアンケート回答内容

2. 実験テーマと実施方法

銀行様

実施 企業名	・株式会社三菱UFJ銀行 ・株式会社りそな銀行 ・株式会社横浜銀行	責任者	・株式会社 FRONTEO
実験 テーマ	現行方法と新方法【AI】を用いた、 <u>業務生産性の比較試験</u>		
対象 業務	投資信託などの金融商品販売時の <u>営業応接記録のチェック業務</u>		
検証 観点	① 保険乗換勧誘時に、<u>解約に伴う不利益事項に関する営業員の説明及びお客様の理解が不十分</u>であると読み取れる記載 ② 営業員からお客様への説明で、<u>元本割れリスクに関する説明・認識が不十分</u>であると読み取れる記載 ③ お客様が<u>過剰な分配金志向を有している</u>と読み取れる記載		
実施 方法	【比較試験】 ・現行方法と新方法【AI】での検出率、業務効率を比較する試験 ・データ準備者とチェック者を分け、複数人のチェック者が同一データに対して異なる方法でチェックを行い、該当データの検出率を測定・比較する ※比較試験実施に耐えうるデータかどうか、事前に教師データの検出精度検証を行う。また1行1観点で実施。		

2. 実験テーマと実施方法

証券様

実施 企業名	・SMBC日興証券 株式会社	責任者	・株式会社 FRONTEO
実験 テーマ	人による確認方法と新方法【AI】を用いた、 <u>業務生産性の比較試験</u>		
対象 業務	顧客との通話時に残されるテキスト変換された <u>通話録音記録のチェック業務</u>		
検証 観点	① お客さまから営業員へ向けられた <u>苦情・不満の声</u>		
実施 方法	【比較試験】 ・人による確認方法と新方法【AI】での検出率、業務効率を比較する試験 ・データ準備者とチェック者を分け、複数人のチェック者が同一データに対して異なる方法でチェックを行い、該当データの検出率を測定・比較する ※比較試験実施に耐えうるデータかどうか、事前に学習済みモデルの検知精度検証を行う		

2. 実験テーマと実施方法 –【参考】 イメージ①全体の流れ

実験前の準備フェーズ

実験フェーズ

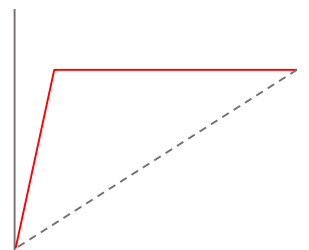
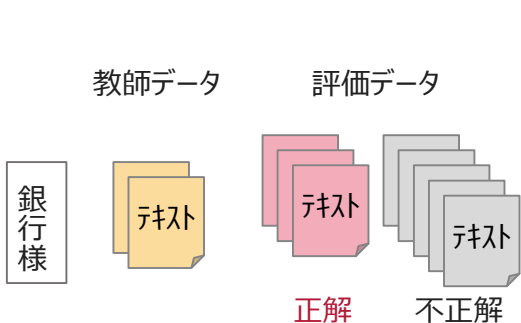
データ準備

学習済みモデル構築/精度確認

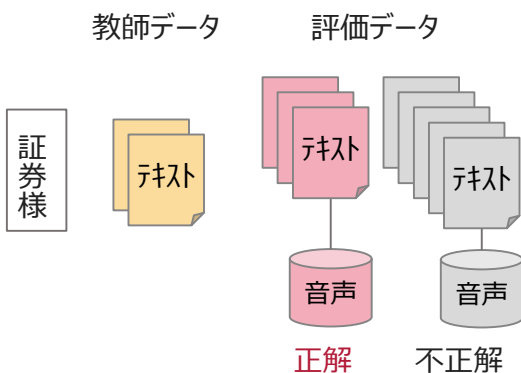
比較試験の実施（対照実験）

□ 正解・不正解データと教師データ準備 □ AI精度確認

□ チェック者とチェック方法・データセットのパターン



- 教師データを用い、比較試験に使用する学習済みモデルを構築する
- 対象観点で実験に使用するに足る精度があることを確認する

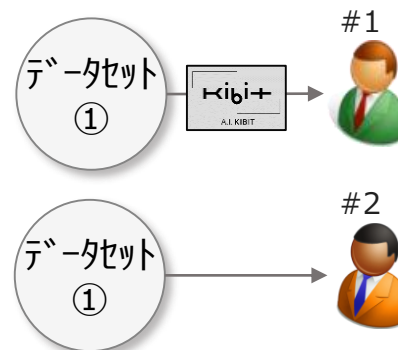


- 教師データに対応するデータセット（2セット）を事前準備
※事前に正解・不正解はチェック済み

		チェック者 #1	チェック者 #2	チェック者 #3
A	現行方法	データセット ①	データセット ②	データセット ①
B	新方法【AI】 KIBIT	データセット ②	データセット ①	データセット ②

- 各データセットが現行・KIBITの両方法でチェックされるとともに、チェック者が同じデータセットをチェックしないよう、割当を行う
※データセット2つ、チェック者（最低2名）

□ チェックの実施



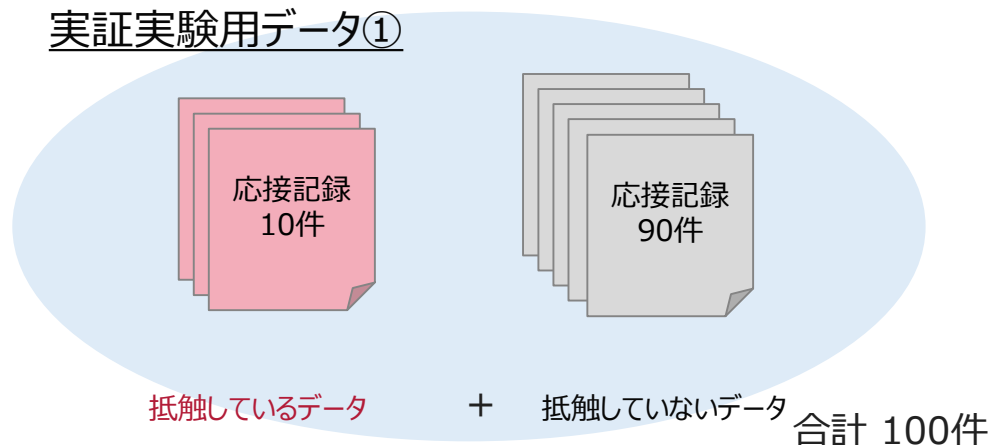
- KIBITスコアを参考に、上位から順にNGデータのチェック
- 音声はNG候補データの詳細確認のため、該当音声データも聞いて最終チェック
※テキスト変換精度低のため
- 音声は音声データそのものを聞いてチェックし、応接記録・メールはテキストデータをチェック
※人による確認方法で行う

2. 実験テーマと実施方法 –【参考】 イメージ②経験の異なる被験者

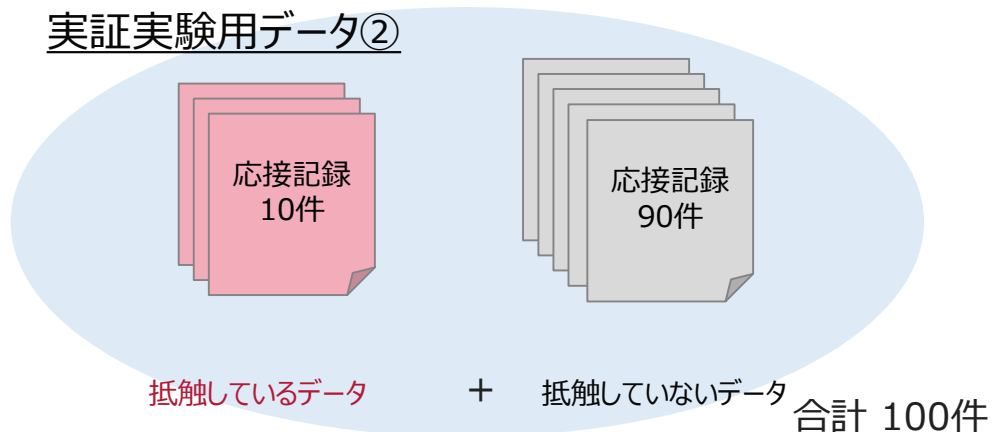
計100件の応接記録の中から、「対象観点の抵触」が読み取れる応接記録を、多く見つけることを目標とする。

ご用意頂くデータ

実証実験用データ①



実証実験用データ②



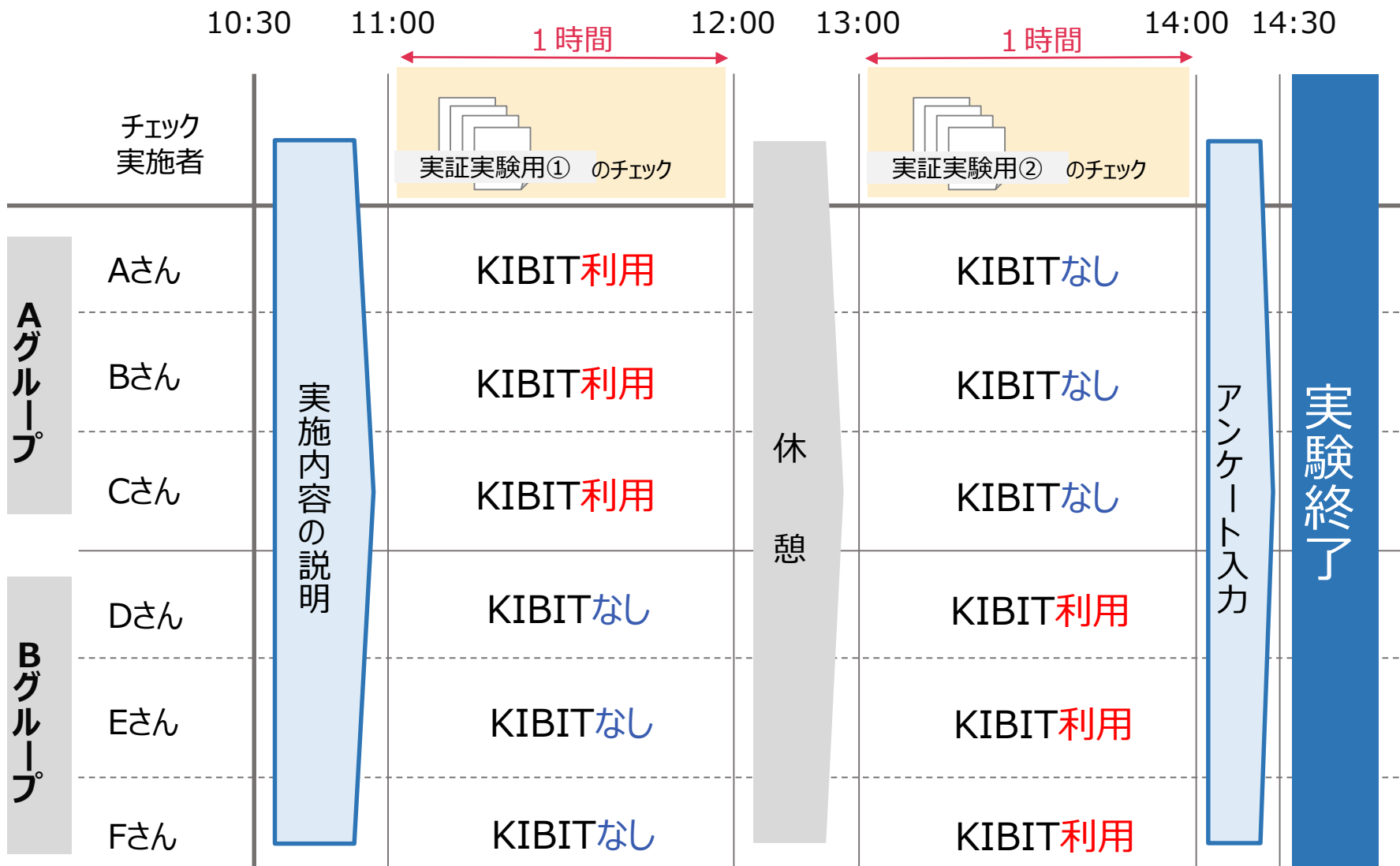
チェック実施者

	Aグループ	Bグループ
業務経験 豊富	Aさん	Dさん
業務経験 やや豊富	Bさん	Eさん
業務経験 普通	Cさん	Fさん

業務経験豊富: チェック業務経験が長く、経験豊富な人
業務経験やや豊富: チェック業務経験がやや豊富の人
業務経験ふつう: チェック業務経験が短い人

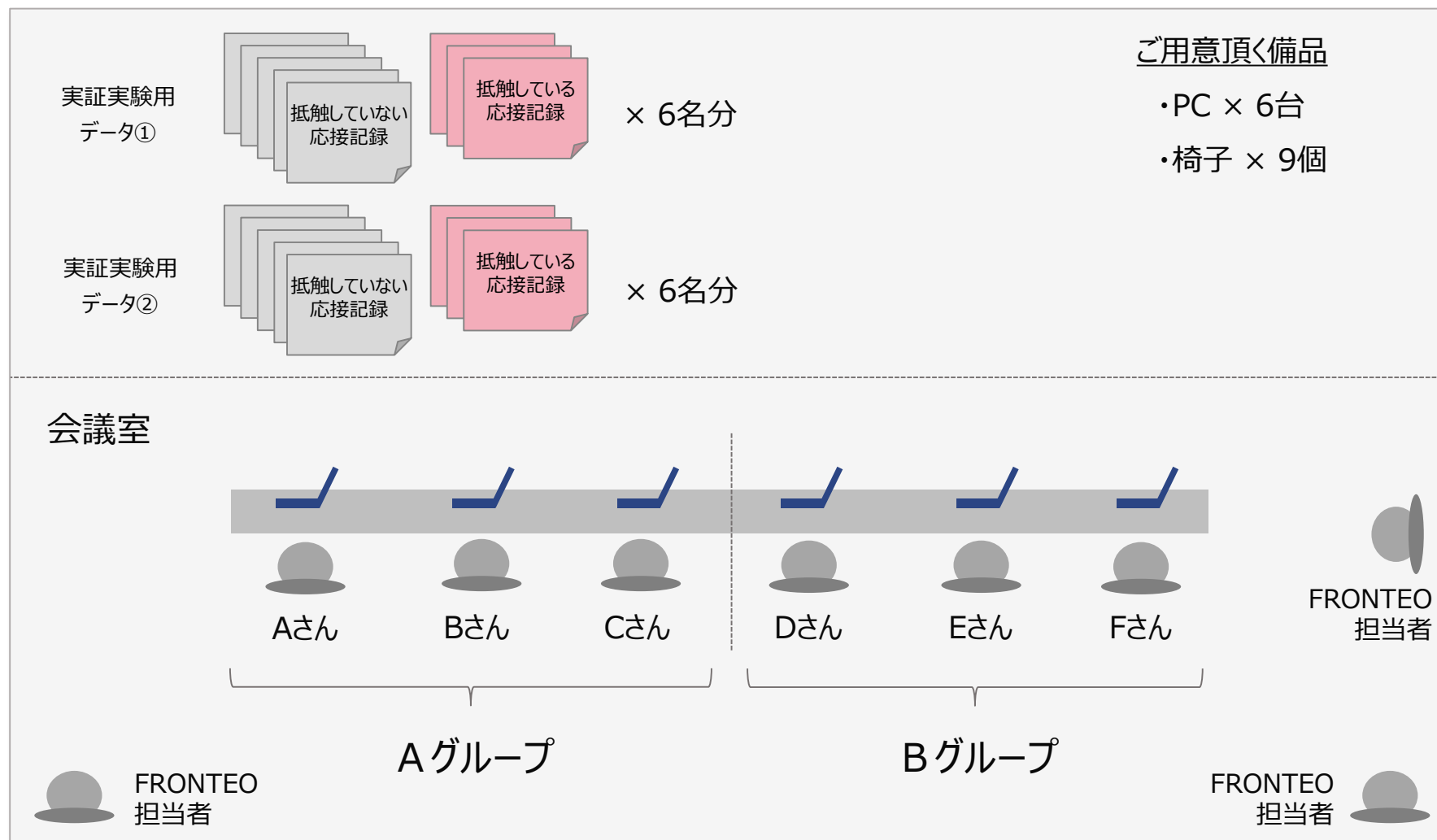
2. 実験テーマと実施方法 –【参考】 イメージ③実験当日の流れ

実験時間は、10:30-14:30を予定。終了後にアンケートを記入。



2. 実験テーマと実施方法 –【参考】イメージ④実験当日のレイアウト

以下のレイアウトで実証実験を実施。



銀行様

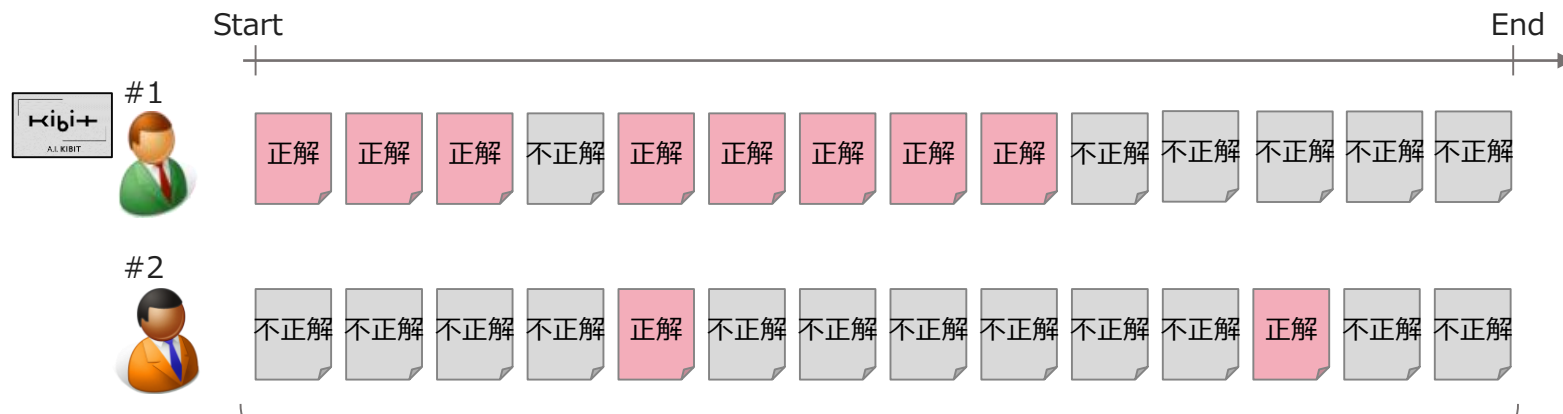
応接記録

(例)

全データ：100件

正解：10件

制限時間：1時間



• 正解数、確認件数、作業終了までの時間（実工数）を測定する

証券様

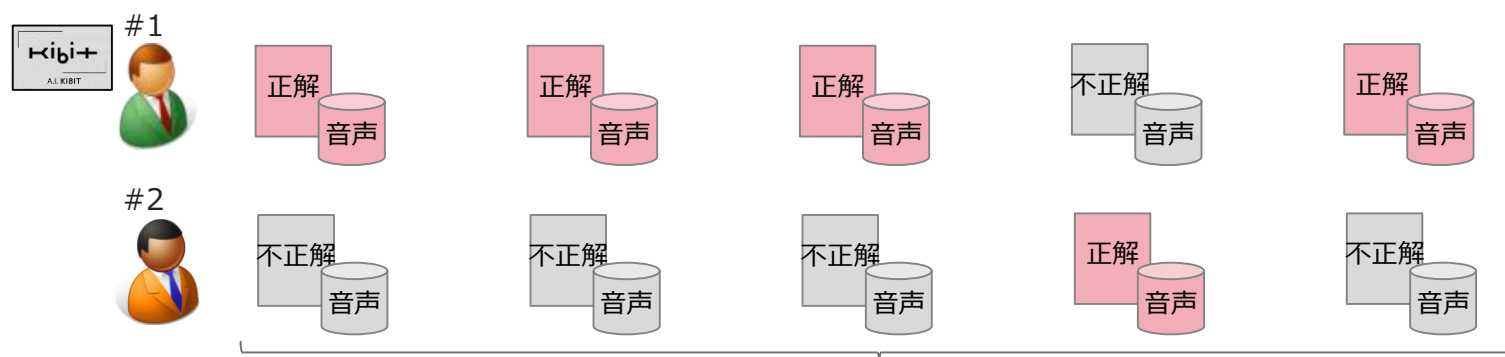
音声通話記録

(例)

全データ：100件

正解：10件

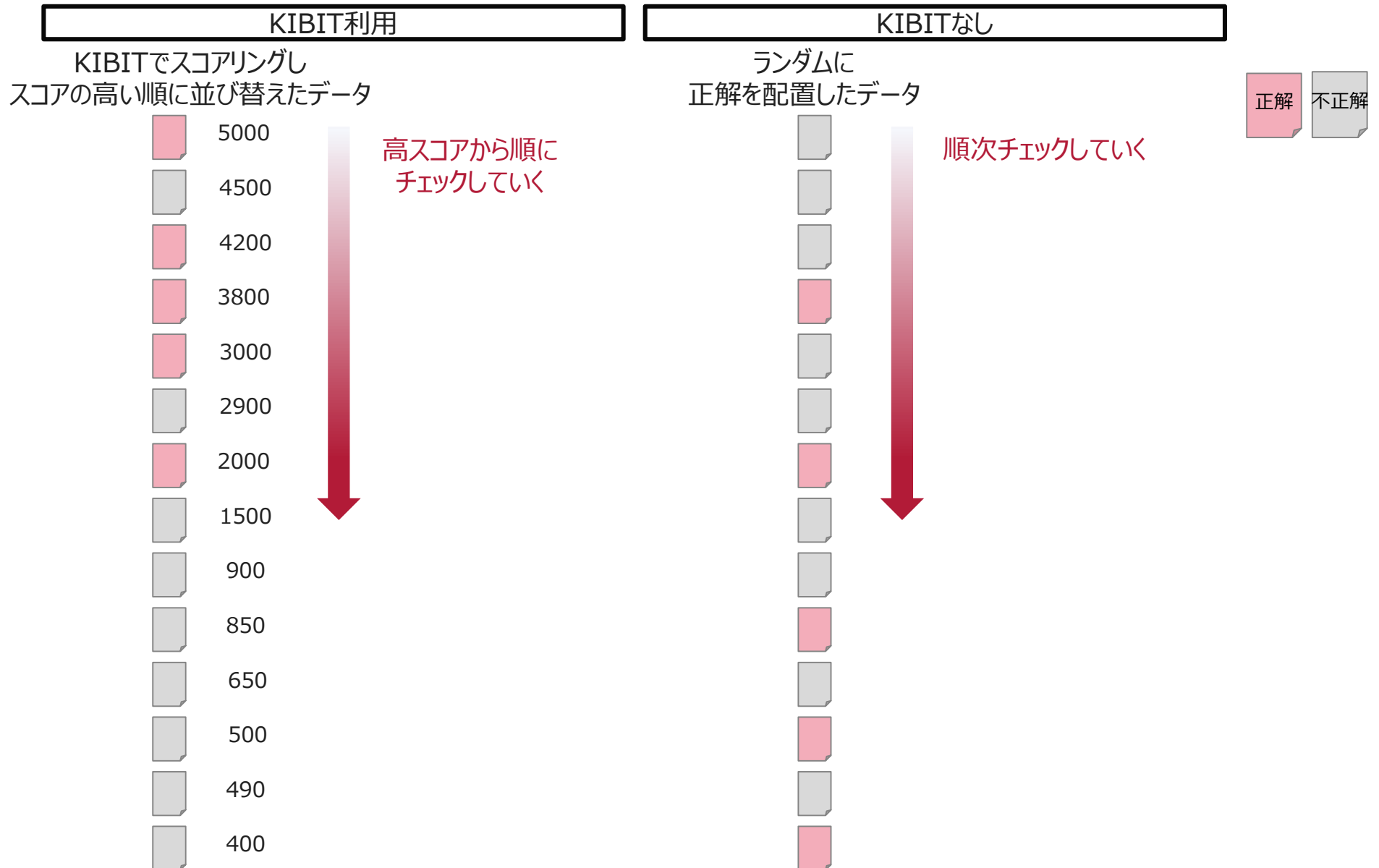
制限時間：1時間



• 正解数、確認件数、作業終了までの時間（実工数）を測定する

個々の実証実験にて、必要な数値を定量的に測定する

2. 実験テーマと実施方法 –【参考】 イメージ⑥チェック方法



1. 本実証実験の概要

1.1 本実証実験の概要および体制図

1.2 実証実験ハブ参加検討に至った背景

1.3 実証実験における目的や目指すところ

2. 実験テーマと実施方法

3. 実験条件

3.1 実験条件詳細

3.2 KIBITによる事前精度検証

4. 実験結果

4.1 結果①応接記録

4.2 結果②お客様との通話記録

5. 結論と考察

5.1 結論と考察

5.2 所感

5.3 監督指針上、確認させていただきたいポイント

5.4 実験参加者のアンケート回答内容

3.1 実験条件詳細

3.1 実験条件詳細

銀行様

観点		観点①	観点②	観点③
観点		保険乗換勧誘時に、解約に伴う不利益事項に関する営業員の説明及びお客様の理解が不十分であると読み取れる記載	営業員からお客様への説明で、元本割れリスクに関する説明・認識が不十分であると読み取れる記載	お客様が過剰な分配金志向を有していると読み取れる記載
対象データ数 -2セット-		応接記録 - データセット① 正解データ5件、不正解データ45件 - データセット② 正解データ4件、不正解データ46件	応接記録 - データセット① 正解データ5件、不正解データ45件 - データセット② 正解5データ件、不正解データ44件	応接記録 - データセット① 正解データ25件、不正解データ175件 - データセット② 正解データ25件、不正解データ175件
被験者		・5人	・6人	・6人
制限時間		・1時間		・2時間
チェック 方法	現行	・ランダムに用意された正解・不正解のデータセットに対して、現行の人による検出方法を用いて、対象観点での抵触有無を確認する。		
	AI	- 対象データをKIBIT（AI）を使い、対象観点にてスコアリングする。 - スコアが付与された対象データを上位から順番に確認し、対象観点での抵触有無を確認する。		

観点①	
観点	お客さまから営業員へ向けられた<u>苦情・不満の声</u>
対象データ数 -2セット-	<u>音声通話記録+テキスト</u> - データセット① 正解データ10件、不正解データ90件 - データセット② 正解データ10件、不正解データ90件
被験者	4人
制限時間	1時間
チェック 方法	人による確認 ランダムに用意された正解・不正解のデータセットに対して、人による確認方法で対象観点である苦情の有無を確認する。 ※まず、テキスト化したデータを確認し、判断に際して詳細確認が必要な場合は、音声データを確認する。
	AI KIBIT（AI）を使い、苦情検知によるスコアリング済みの対象データ（テキスト化ファイル）を上位から順番に確認し、苦情有無を確認する。 ※まず、テキスト化したデータを確認し、判断に際して詳細確認が必要な場合は、音声データを確認する。

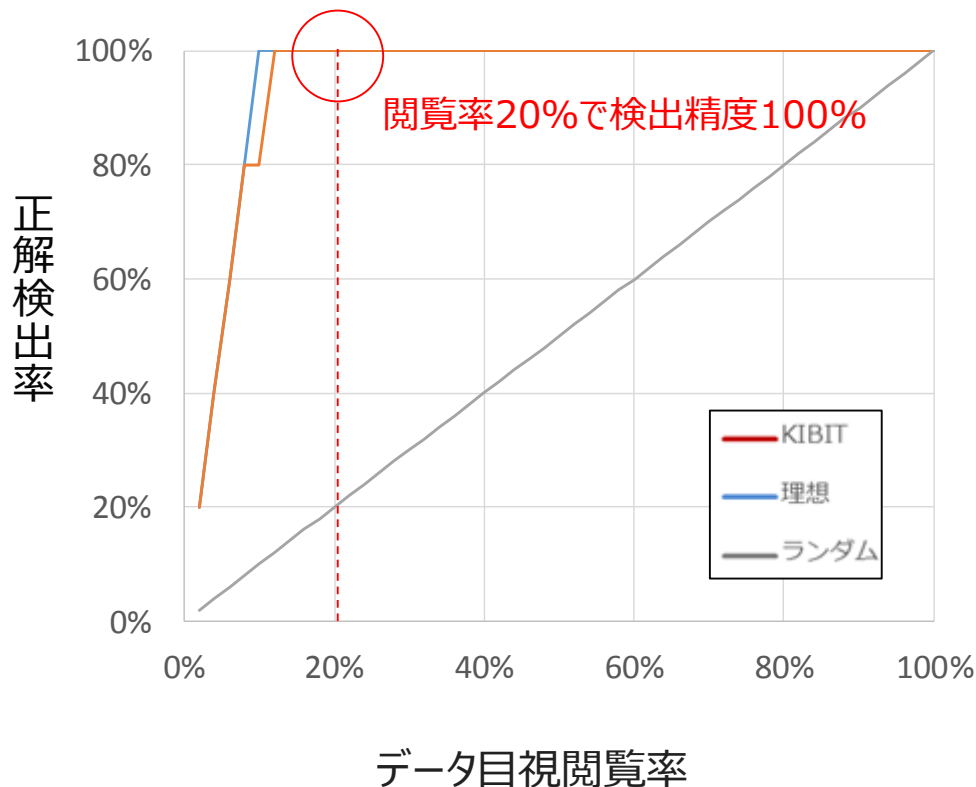
3.2 KIBITの学習モデルの事前精度検証

3.2 事前精度検証 - 観点① 保険乗換勧誘時に、解約に伴う不利益事項に関する 営業員の説明及びお客様の理解が不十分であると読み取れる記載

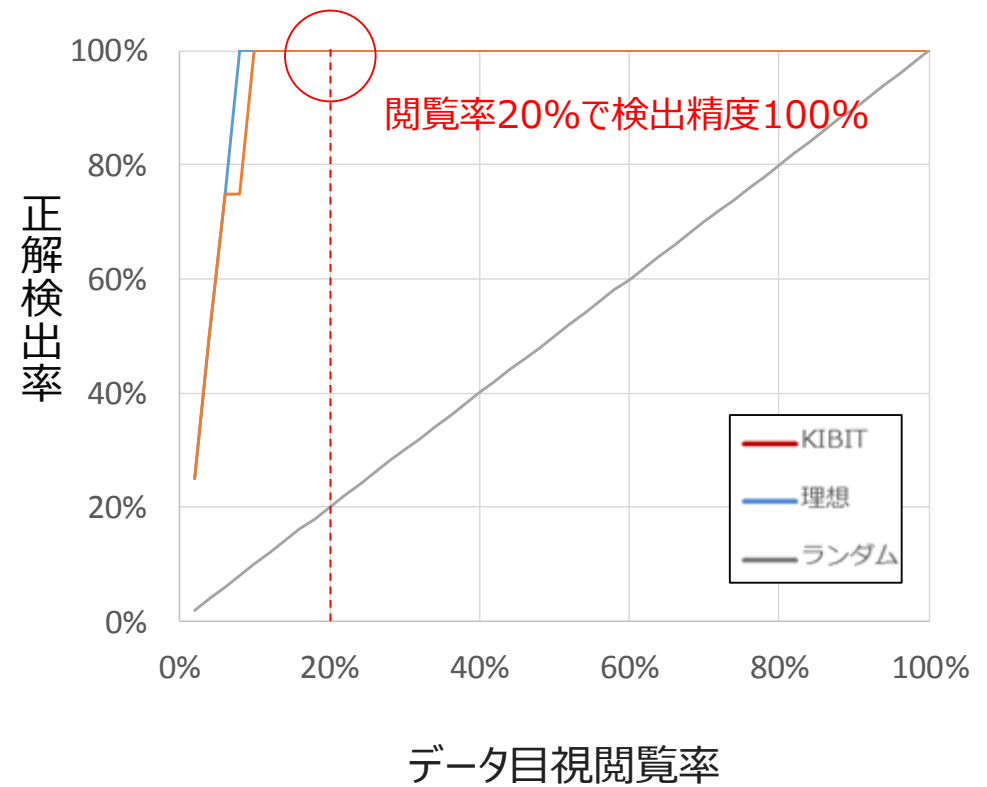
銀行様

データセット①及び②において、閲覧率20%のときの正解検出率が100%と、正解データの抽出実証実験で使用するに十分な精度が得られた。

データセット①に対する精度検証結果



データセット②に対する精度検証結果

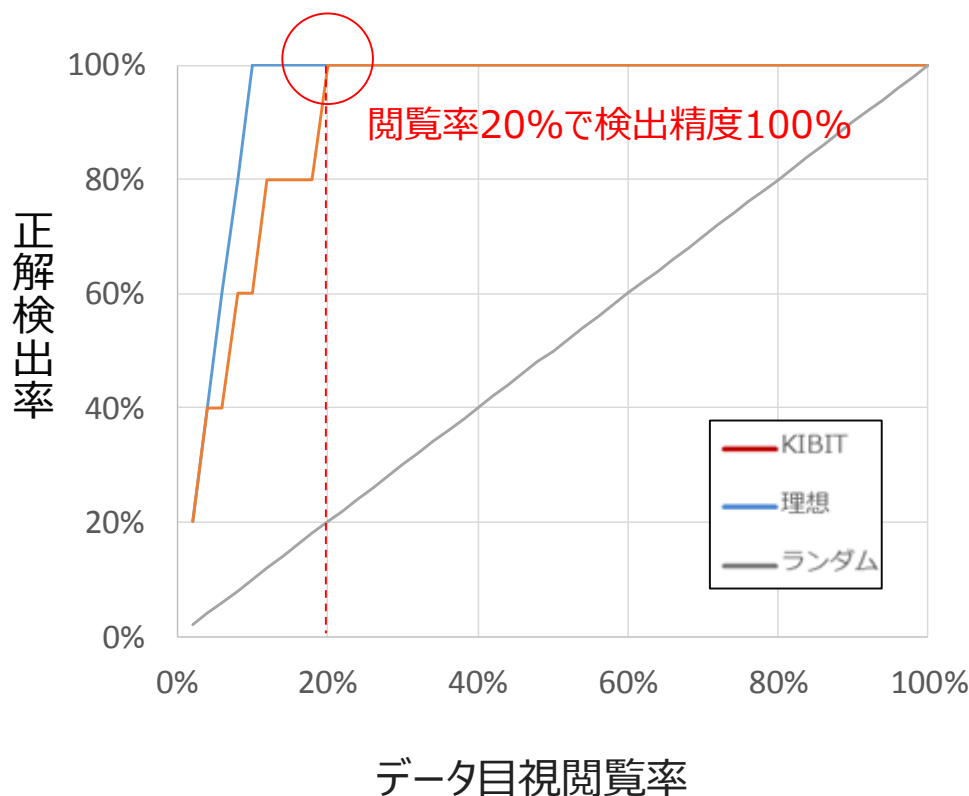


3.2 事前精度検証 - 観点② 営業員からお客様への説明で、 元本割れリスクに関する説明・認識が不十分であると読み取れる記載

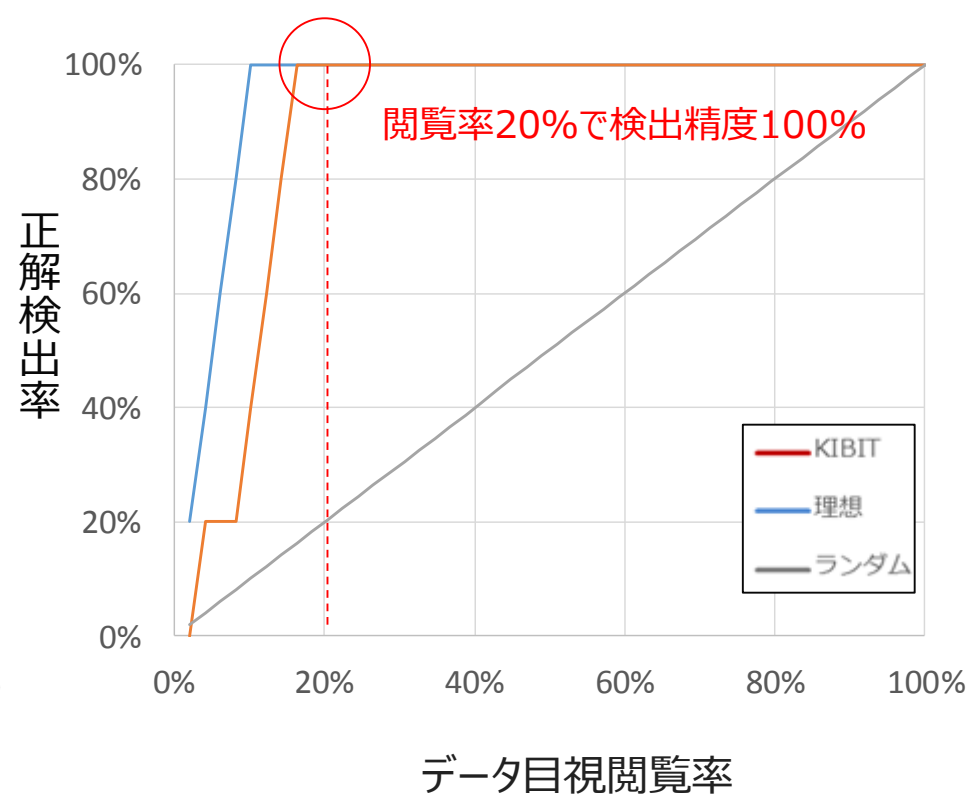
銀行様

データセット①及び②において、閲覧率20%のときの正解検出率が100%と、正解データの抽出実証実験で使用するに十分な精度が得られた。

データセット①に対する精度検証結果



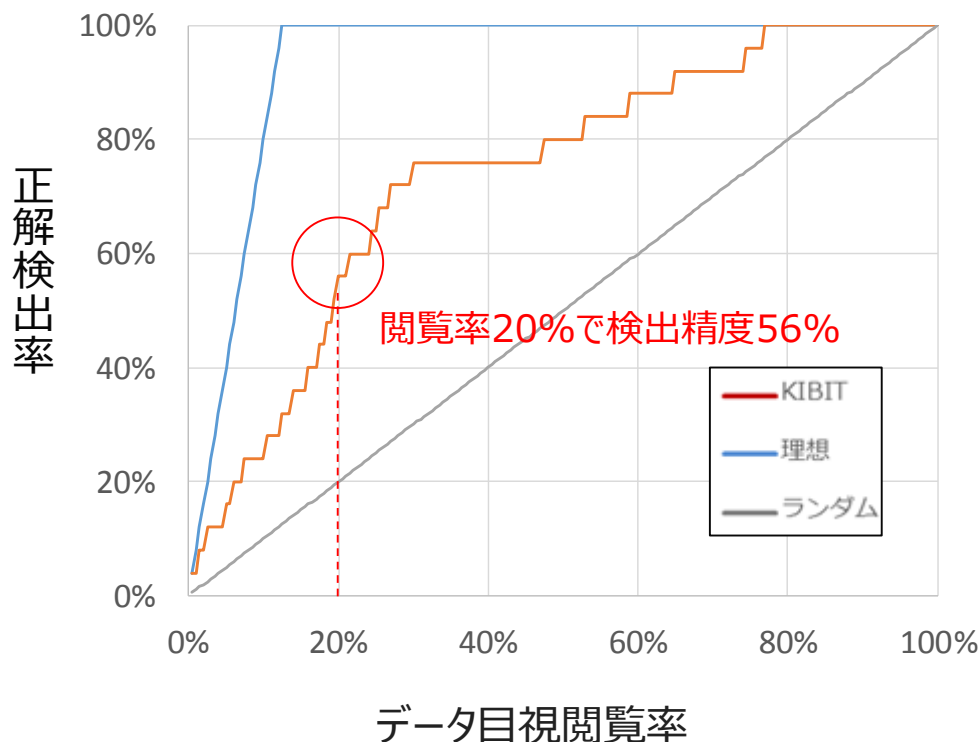
データセット②に対する精度検証結果



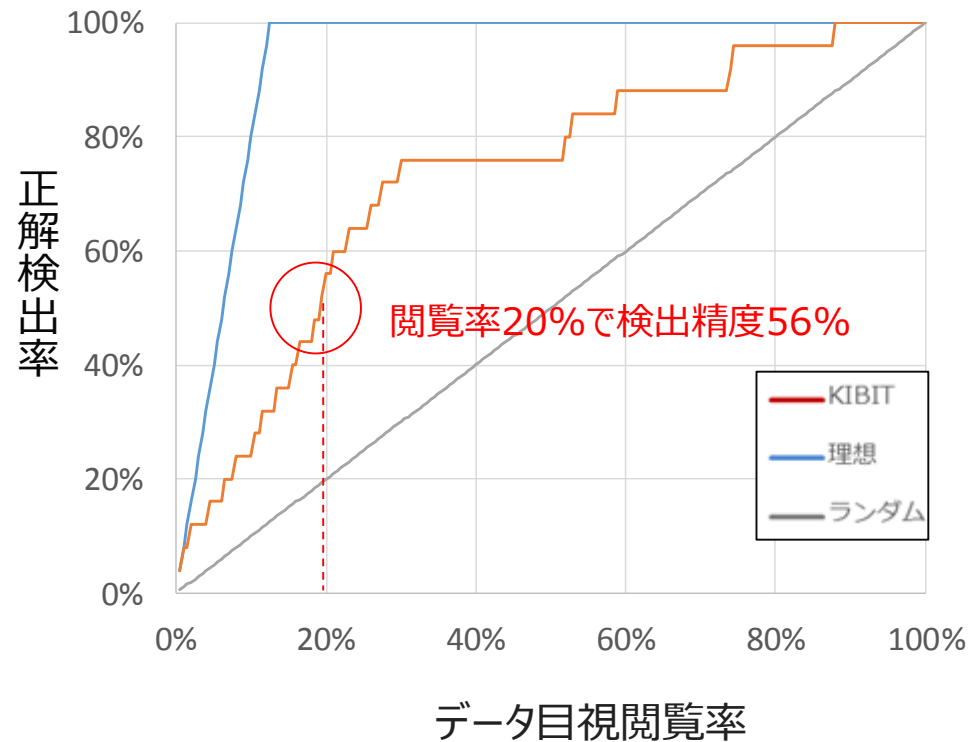
銀行様

データセット①及び②において、閲覧率20%のときの正解検出率が50%以上と、正解データの抽出実証実験で使うことが可能な精度が得られた。

データセット①に対する精度検証結果



データセット②に対する精度検証結果

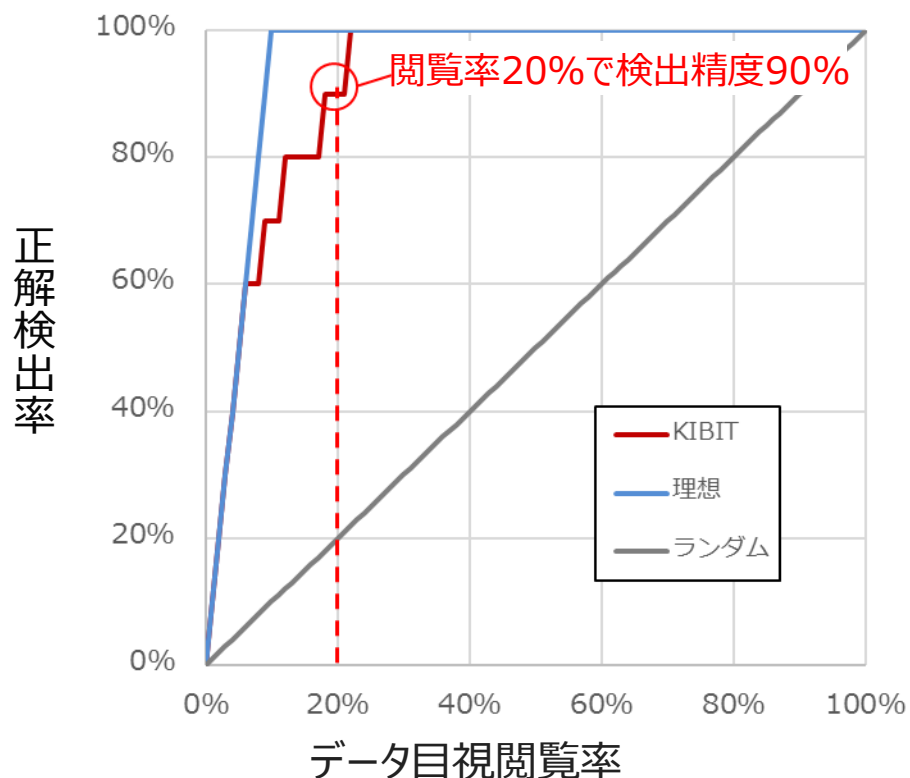


3.2 事前精度検証 - 観点①（音声）お客さまから従業員へ向けられた苦情・不満の声

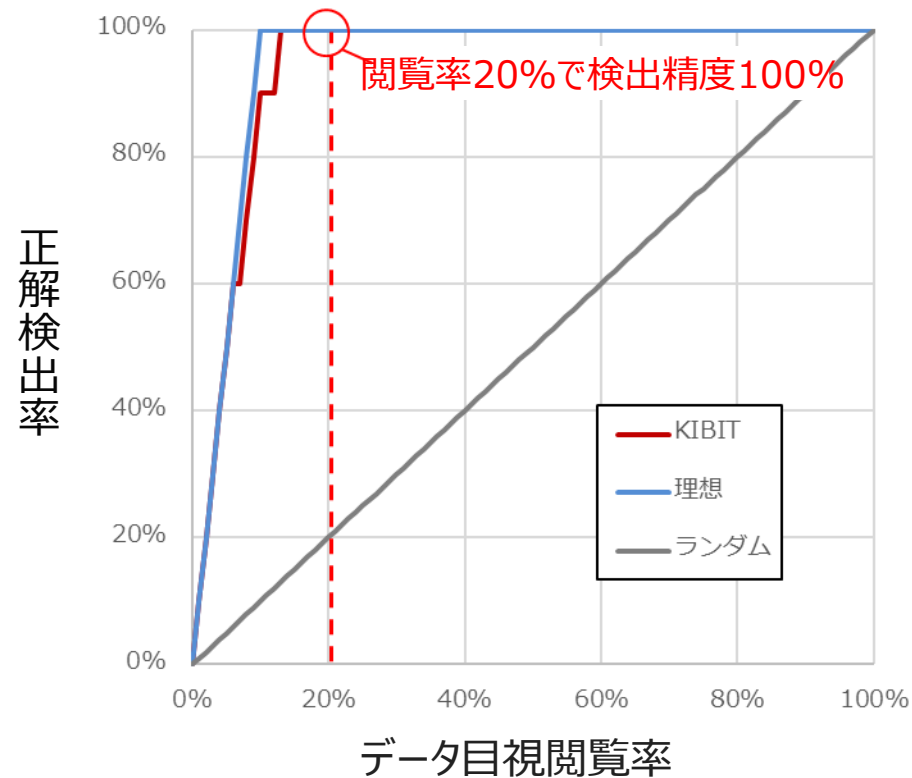
証券様

データセット①及び②において、閲覧率20%のときの正解検出率が90%以上と、正解データの抽出実証実験で使用するに十分な精度が得られた。

データセット①に対する精度検証結果



データセット②に対する精度検証結果



1. 本実証実験の概要
 - 1.1 本実証実験の概要および体制図
 - 1.2 実証実験ハブ参加検討に至った背景
 - 1.3 実証実験における目的や目指すところ
2. 実験テーマと実施方法
3. 実験条件
 - 3.1 実験条件詳細
 - 3.2 KIBITによる事前精度検証
4. 実験結果
 - 4.1 結果①応接記録
 - 4.2 結果②お客様との通話記録
5. 結論と考察
 - 5.1 結論と考察
 - 5.2 所感
 - 5.3 監督指針上、確認させていただきたいポイント
 - 5.4 実験参加者のアンケート回答内容

4. 実験結果

本実験では、比較実験の測定結果を基に、以下の指標を測定した。

	測定指標	定義	測定目的
精度	正解検出率	正解検出件数 ÷ 正解数	• 現行方法、新方法【AI】の両業務の、精度の差異を明確にする
	時間当り正解検出件数	正解検出件数 ÷ 時間	• 現行方法、新方法【AI】の両業務の、生産性の差異を明確にする
効率化 (生産性・省力化)	生産性向上率	新方法【AI】の時間当り正解検出件数 ÷ 現行方法の時間当り正解検出件数	• 現行方法、新方法【AI】の両業務の、生産性の差異を明確にする
	分散比	時間当たり正解検出件数の分散の比率 * 新方法【AI】の分散 ÷ 現行方法の分散	• 現行方法、新方法【AI】の両業務の、人のスキル（属人スキル）による作業バラつきも踏まえた標準化の傾向を明確にする
標準化 (作業品質)	適合率・再現率分布 (+F値)	* 適合率・再現率分布 = 適合率と再現率からグラフ化 * F値 = $2 * \text{適合率} * \text{再現率} \div (\text{適合率} + \text{再現率})$ ※ 両指標の調和平均を取った評価指標	• 現行方法の一連の検出プロセスと、KIBITを活用した一連の検出プロセスを各々システムと見なし、システムの検出能力の差異を明確にする • 曲線が上に位置する程、優れた検出能力を持つ
高度化 (検出能力)	適合率、再現率とは	* 適合率 = 正解検出件数 ÷ 確認件数 * 再現率 = 正解検出件数 ÷ 用意した正解数	• 適合率：現行方法と新方法【AI】の両業務の実務ベースの検出精度を明確にする • 再現率：現行方法と新方法【AI】の両業務の見つけるべきものを見つけれられた割合を明確にする

* 正解 = 業務上の見つけたいテキスト情報のこと

銀行様

4.1 結果①応接記録

4.1 結果①応接記録 【精度】検出率とチェック時間

銀行様

AIを用いた新方法では、③全件チェック完了時（推定）の正解検出率は現行方法と同等以上、また検出するまでの時間も現行方法に比べて約38%短かった。

②実験終了時での正解検出率も、現行方法と比べてAIを用いた新方法が高かった。

■ 各観点での 被験者の一人あたり平均 実施時間・正解検出率

観点①
③

・応接記録における
観点①～③の平均から
算出

①実験終了時

新方法【AI】(実測値)

時間
(分)

74分

正解検出率
(%)

65%

②実験終了時

現行方法(実測値)

時間
(分)

74分

正解検出率
(%)

44%

③全件チェック 完了時(推定) * 1

現行方法(推定値)

時間
(分)

118分

正解検出率
(%)

61%

新方法【AI】は現行方法と同等以上の検出率

新方法【AI】は現行方法よりも38%早く終了

※正解検出率=正解検出件数÷用意した正解数

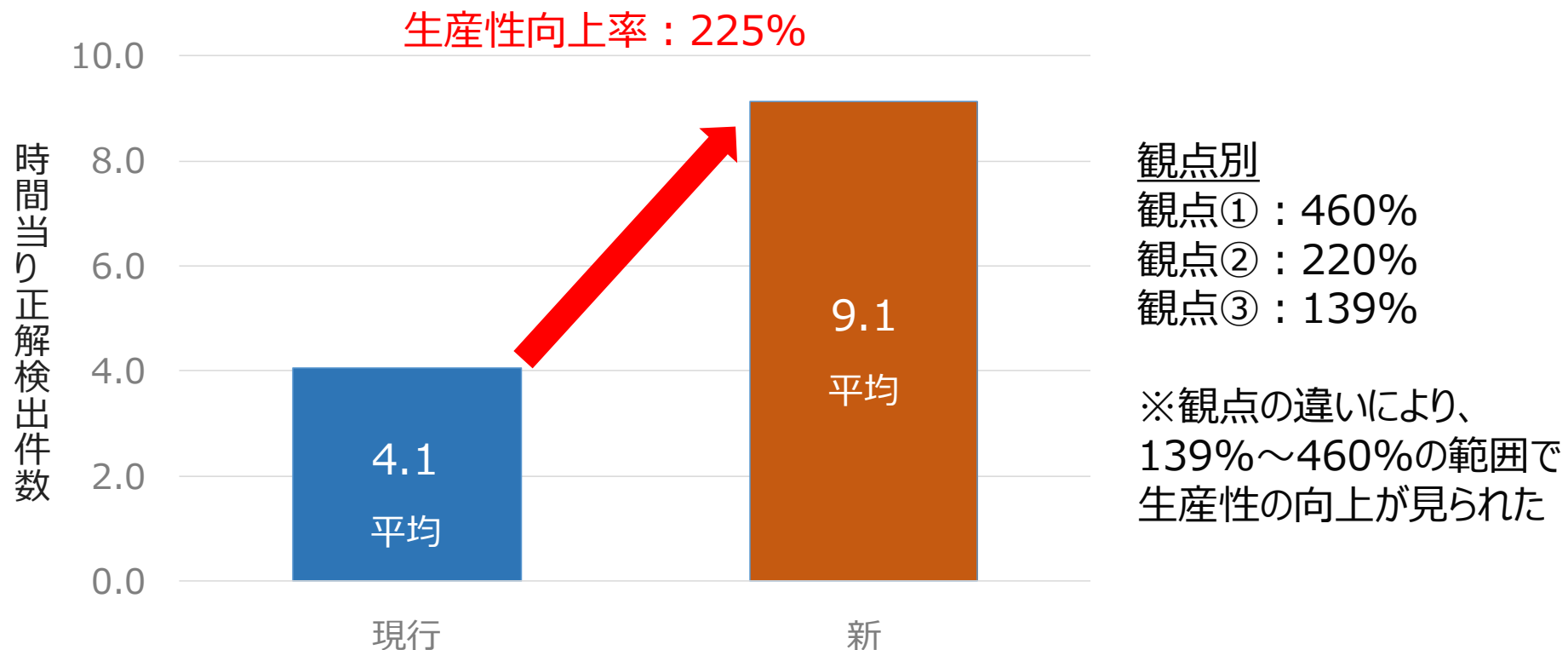
* 1 実験中と同精度・速度でチェックが完了したと仮定し、推計した結果

4.1 結果①応接記録【効率化】時間あたり正解検出件数と生産性向上率

銀行様

観点①～③の各観点にて算出した時間あたり正解検出件数の平均は
(現行) 4.1件/時間⇒ 新方法【AI】9.1件/時間と、大幅な改善がみられた。
【生産性向上率：225%】

時間あたり正解検出件数と生産性向上率（観点①②③、被験者数=17）



※時間あたり正解検出件数=正解検出件数÷時間

※生産性向上率=新方法【AI】の時間あたり正解検出件数÷現行方法の時間あたり正解検出件数

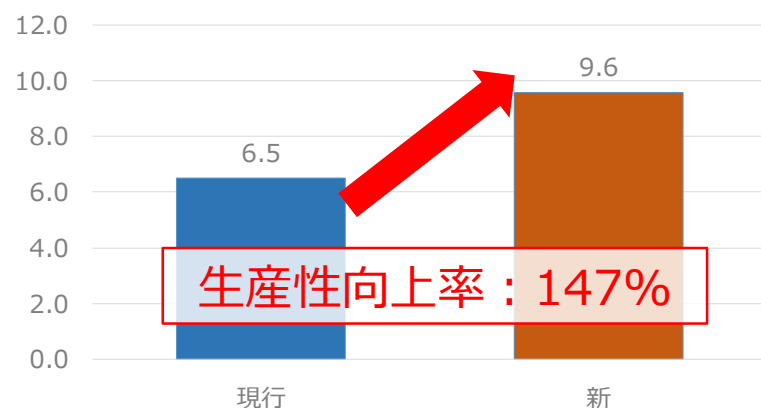
4.1 結果①応接記録【標準化】経験別による生産性向上率

補足①

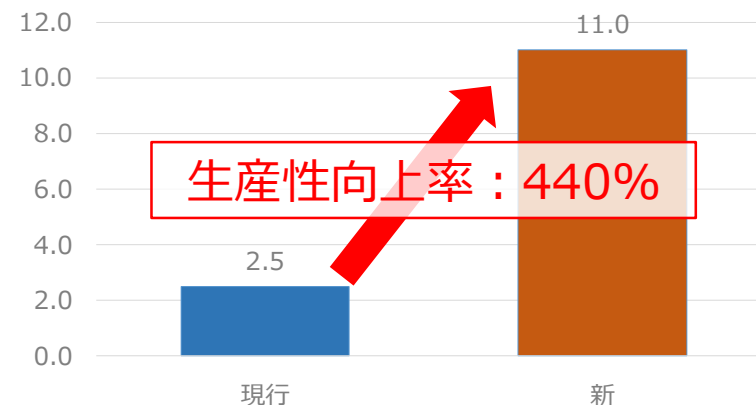
銀行様

※観点①～③全体での被験者の経験別生産性向上率

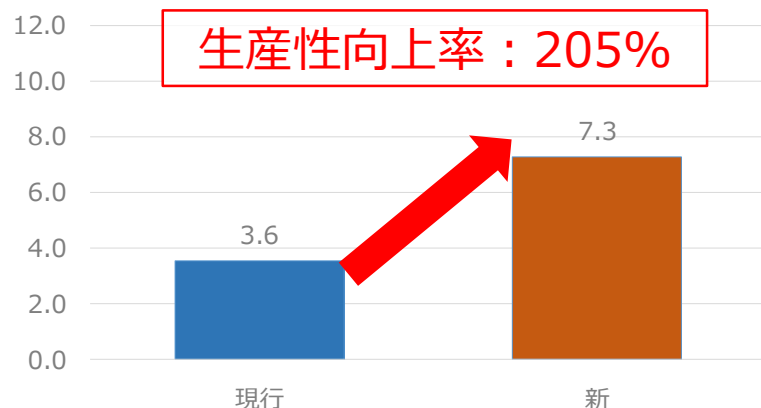
経験豊富（被験者数=6）



経験やや豊富（被験者数=4）



経験普通（被験者数=7）



経験が短い被験者の方が、
AIを活用して効率化が高まる傾向がみられた

*経験普通の新方法では
1名の実験データがエラーのため除外し、6名で集計

※時間当り正解検出件数=正解検出件数÷時間

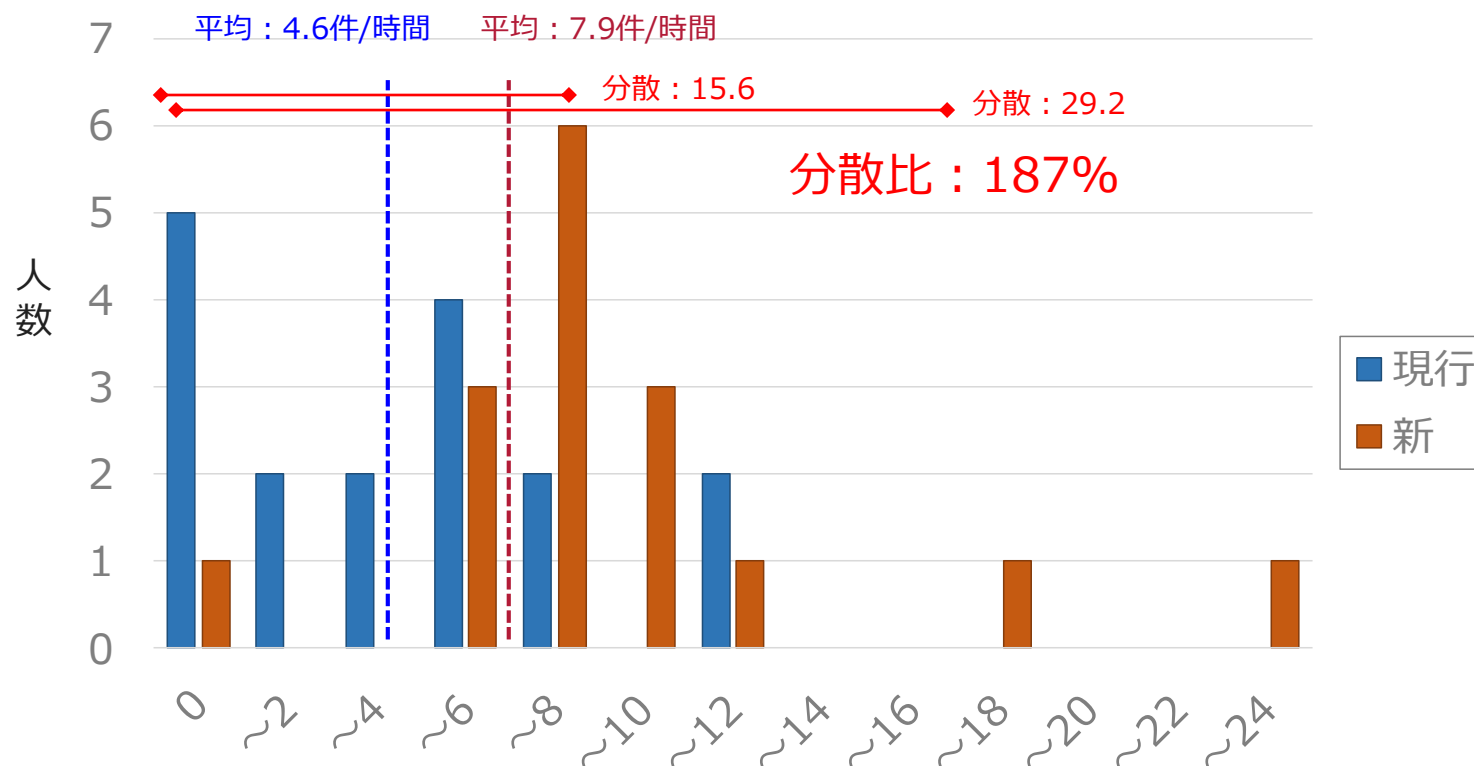
※生産性向上率=新方法【AI】の時間当り正解検出件数÷現行方法の時間当り正解検出件数

4.1 結果①応接記録【標準化】分散

銀行様

観点①～③を含む全体では、時間当り正解抽出件数の分散は、（現行）15.6、（新方法AI）29.2となり、分散比は拡大した。複数観点が混ざったためと考えられる（補足②参照）。

時間当り正解検出件数のヒストグラムと分散比（被験者数=17）



*新方法では1名の実験データがエラーのため除外し、16名で集計

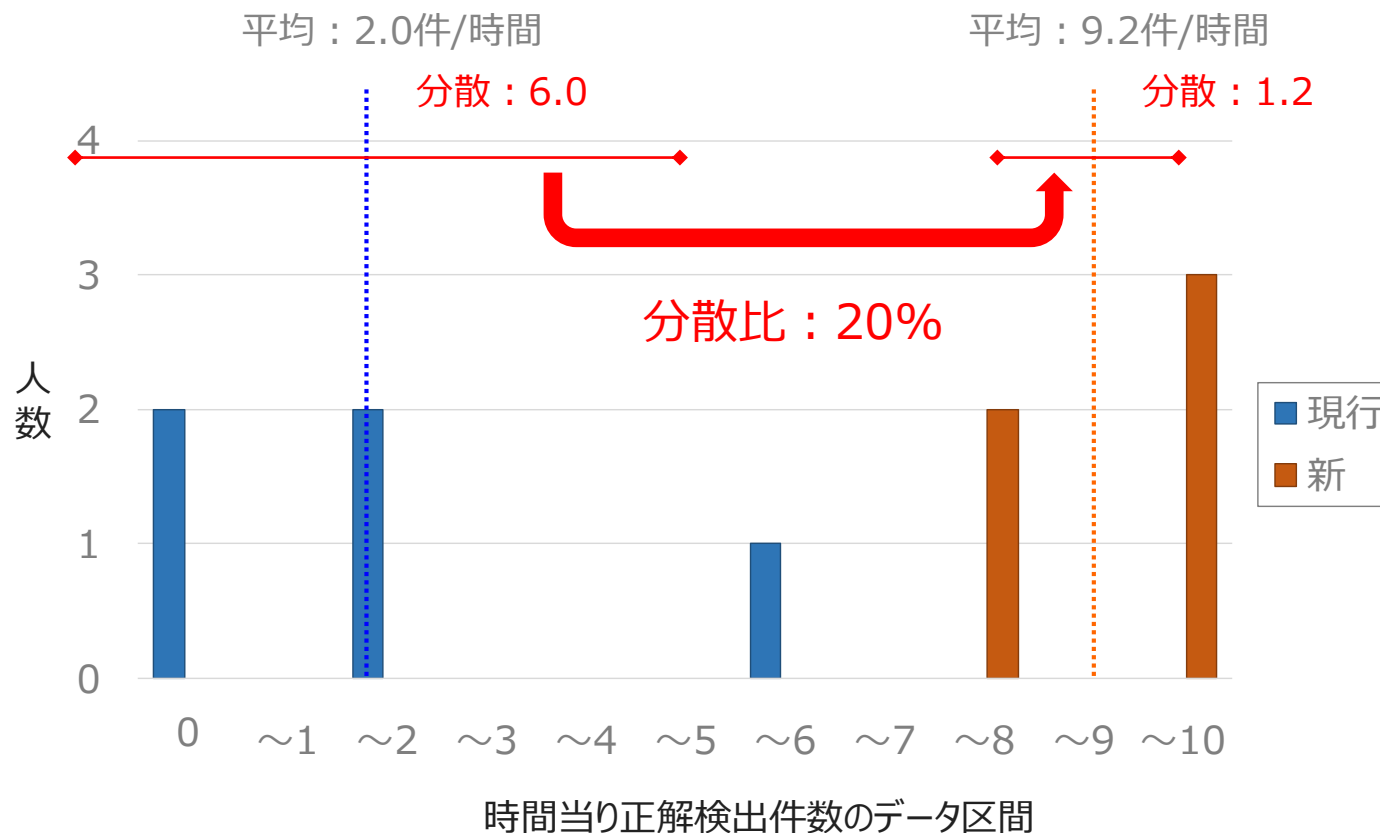
時間当り正解検出件数のデータ区間
分散比 = 新方法【AI】の分散 ÷ 現行方法の分散

銀行様

観点①における分散比

【観点①】

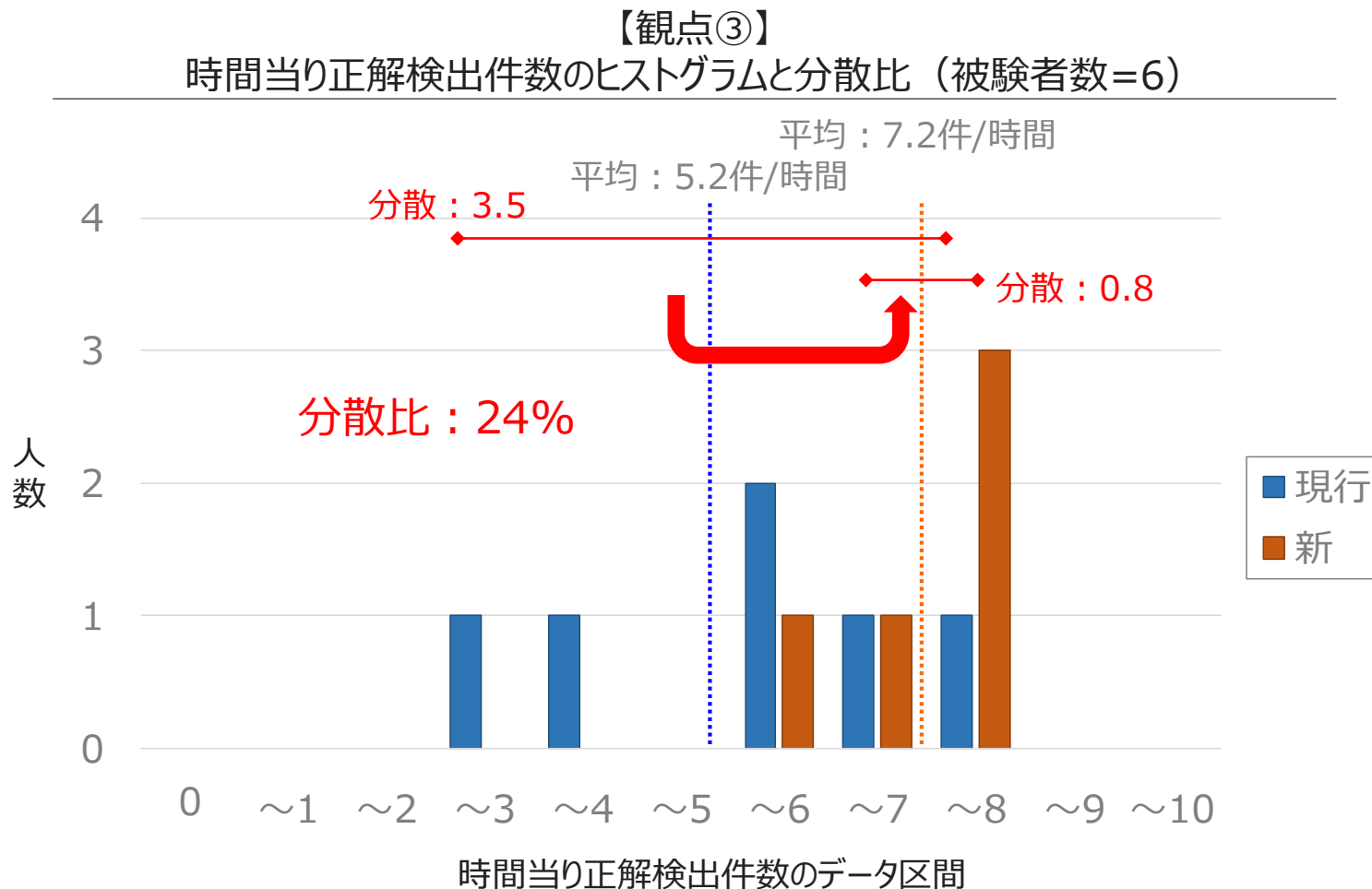
時間当り正解検出件数のヒストグラムと分散比（被験者数=5）



観点別で見ると、ばらつきは少なくなり、標準化の効果が確認された

分散比 = 新方法【AI】の分散 ÷ 現行方法の分散

銀行様 観点③における分散比



観点別で見ると、ばらつきは少なくなり、標準化の効果が確認された

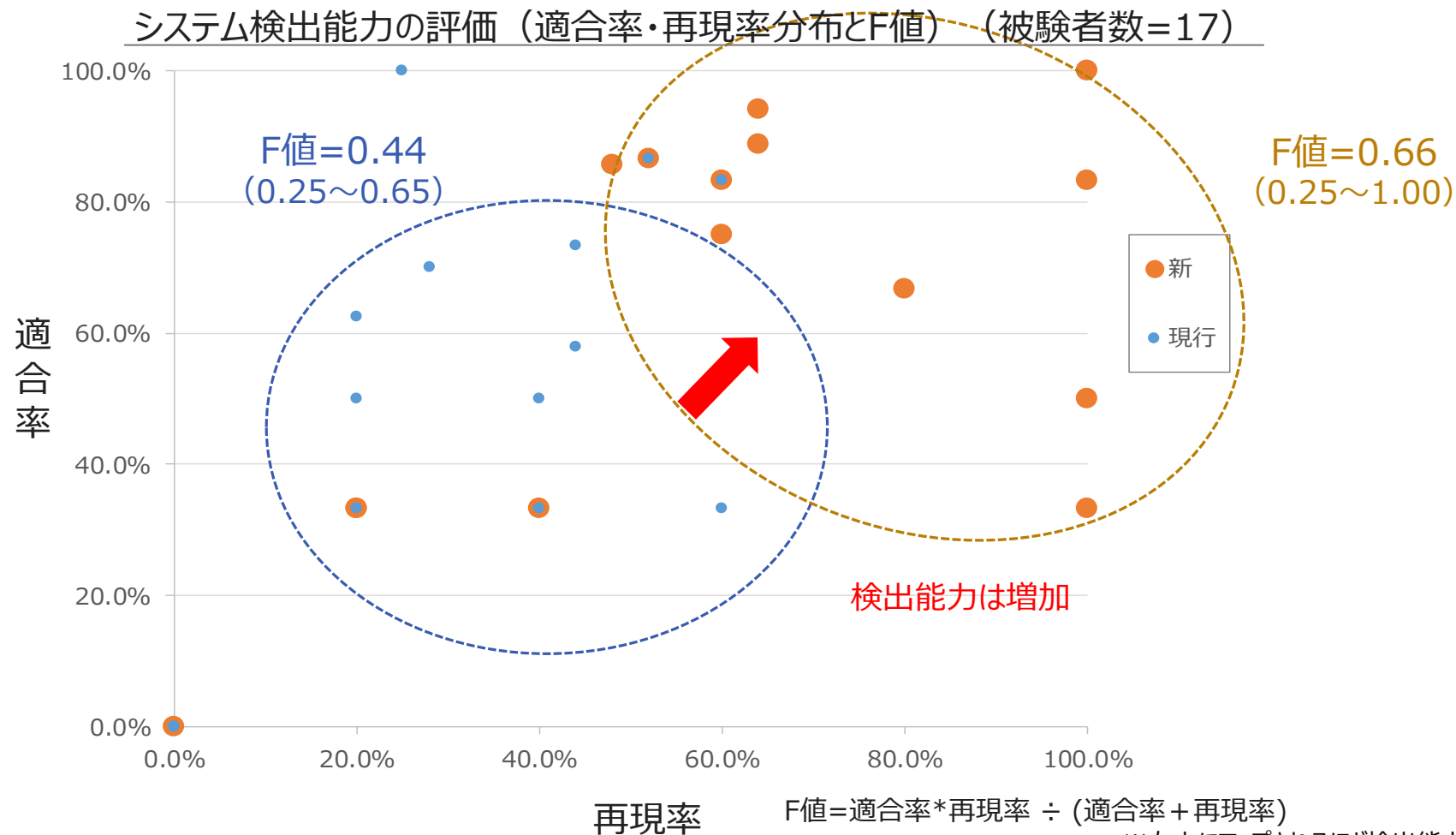
$$\text{分散比} = \text{新方法【AI】の分散} \div \text{現行方法の分散}$$

*新方法では1名の実験データがエラーのため除外し、5名で集計

4.1 結果①応接記録【高度化】適合率・再現率（システム検出能力）

銀行様

適合率と再現率の分布より、現行方法より新方法【AI】の方が、システム検出能力は高まった傾向が確認された。
(F値で0.44⇒0.66 (150%) へ増加)



証券様

4.2 結果②お客様との通話記録

4.2 結果②お客様との通話記録 【精度】検出率とチェック時間

証券様

AIを用いた新方法では、③全件チェック完了時（推定）の正解検出率は現行方法と同等以上、また検出するまでの時間も現行方法に比べて約55%短かった。

②実験終了時での正解検出率も、現行方法と比べてAIを用いた新方法が高かった。

■ 各観点での 被験者の一人あたり平均 実施時間・正解検出率

①実験終了時

②実験終了時

③全件チェック 完了時(推定) * 1

新方法AI (実測値)	
時間 (分)	正解検出率 (%)

現行方法(実測値)	
時間 (分)	正解検出率 (%)

現行方法(推定値)	
時間 (分)	正解検出率 (%)

60分

75%

60分

40%

131分

72%

新方法【AI】は現行方法と同等以上の検出率

新方法【AI】は現行方法よりも55%早く終了

観点①

・お客さまから営業員へ向けられた苦情・不満の声

※正解検出率=正解検出件数÷用意した正解数

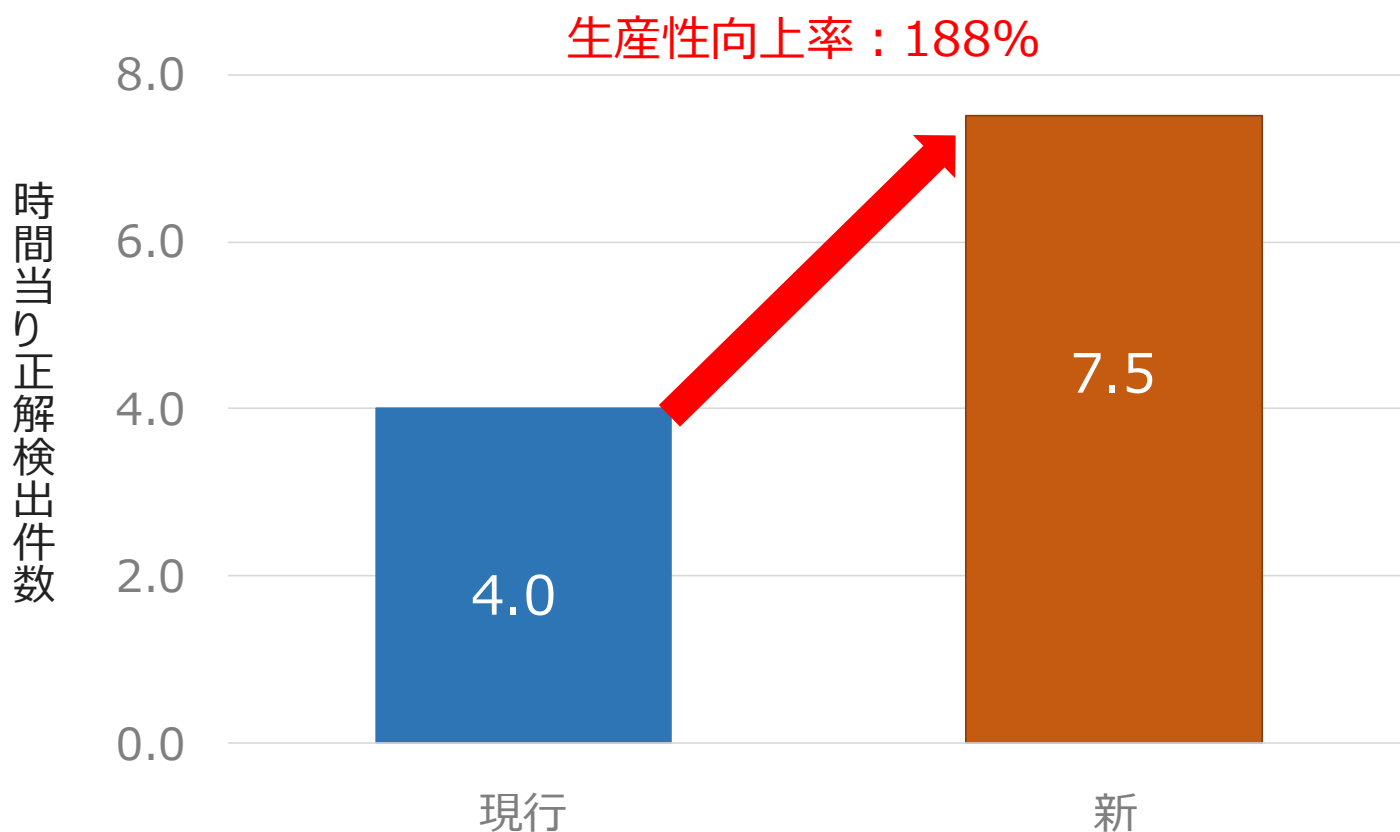
* 1 実験中と同精度・速度でチェックが完了したと仮定し、推計した結果

4.2 結果②お客様との通話記録 【効率化】時間あたり正解検出件数と生産性向上率

証券様

全体では時間当り正解検出件数が（現行）4.0件/時間⇒（新方法AI）7.5件/時間と、改善がみられた。
（生産性向上率：188%）

時間当り正解検出件数と生産性向上率（現行方法：被験者数=4、新方法【AI】：被験者数=2）



※時間当り正解検出件数=正解検出件数÷時間

※生産性向上率=新方法【AI】の時間当り正解検出件数÷現行方法の時間当り正解検出件数

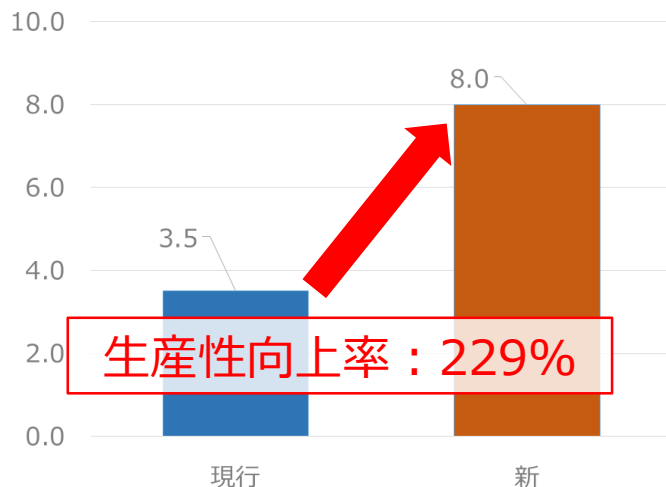
4.2 結果②お客様との通話記録【標準化】経験別による生産性向上率

証券様

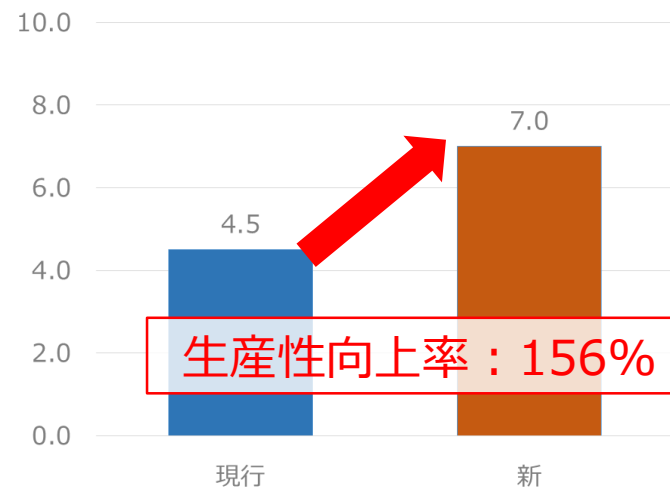
※被験者の経験別生産性向上率

時間
当り
正解
検出
件数

経験豊富（被験者数=2）



経験普通（被験者数=2）



経験が長い被験者の方が、AI活用による生産性向上率が高い傾向がみられた

※時間当り正解検出件数=正解検出件数÷時間

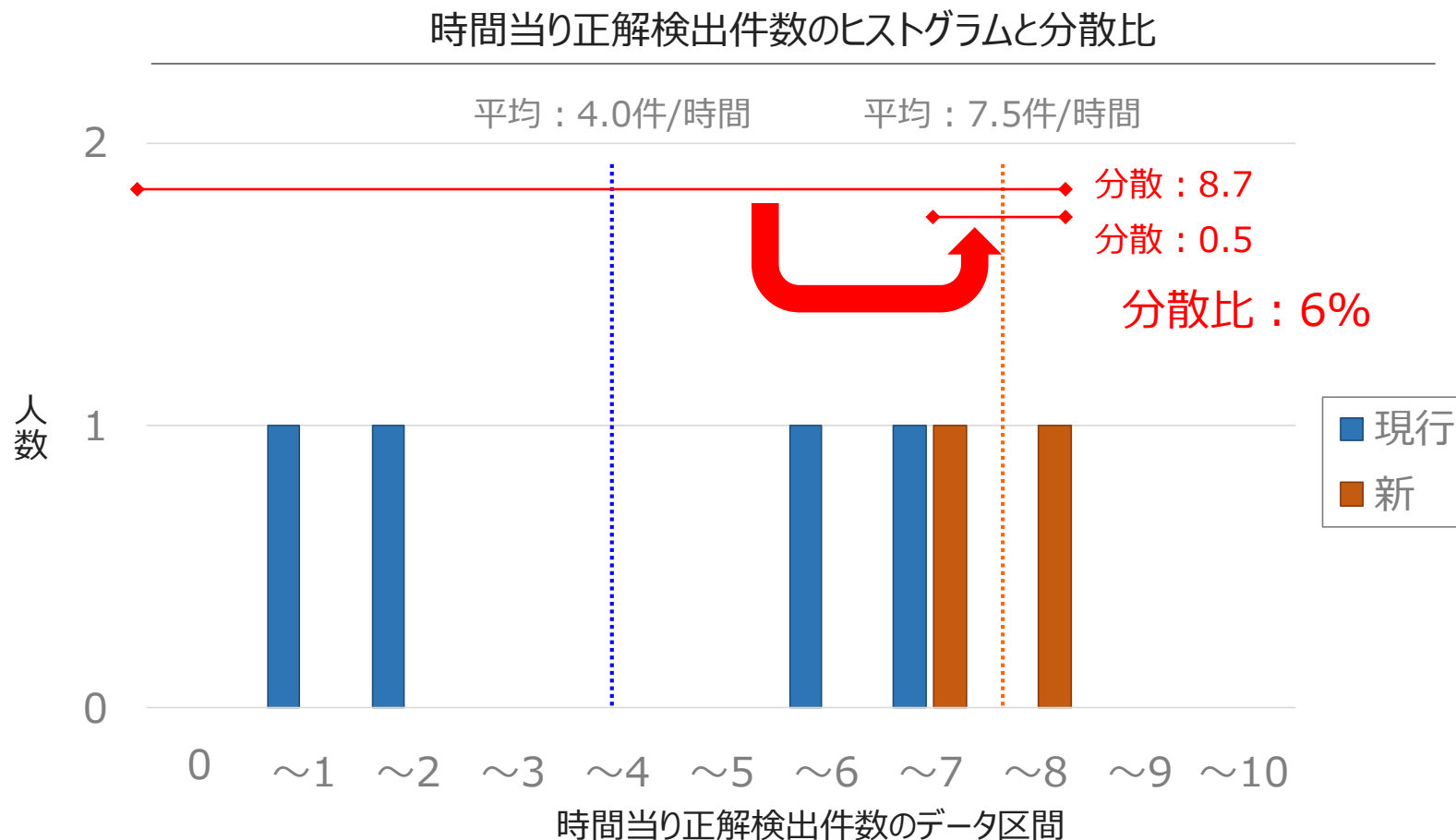
※生産性向上率=新方法【AI】の時間当り正解検出件数÷現行方法の時間当り正解検出件数

*新方法では経験豊富/普通共に
1名の実験データがエラーのため除外し、1名で集計

4.2 結果②お客様との通話記録 【標準化】分散比

証券様

時間当り正解抽出件数の分散は、現行方法 8.7、新方法【AI】 0.5となった。分散比からばらつき差異を算出すると、分散比は6%となり、標準化の効果が確認された。



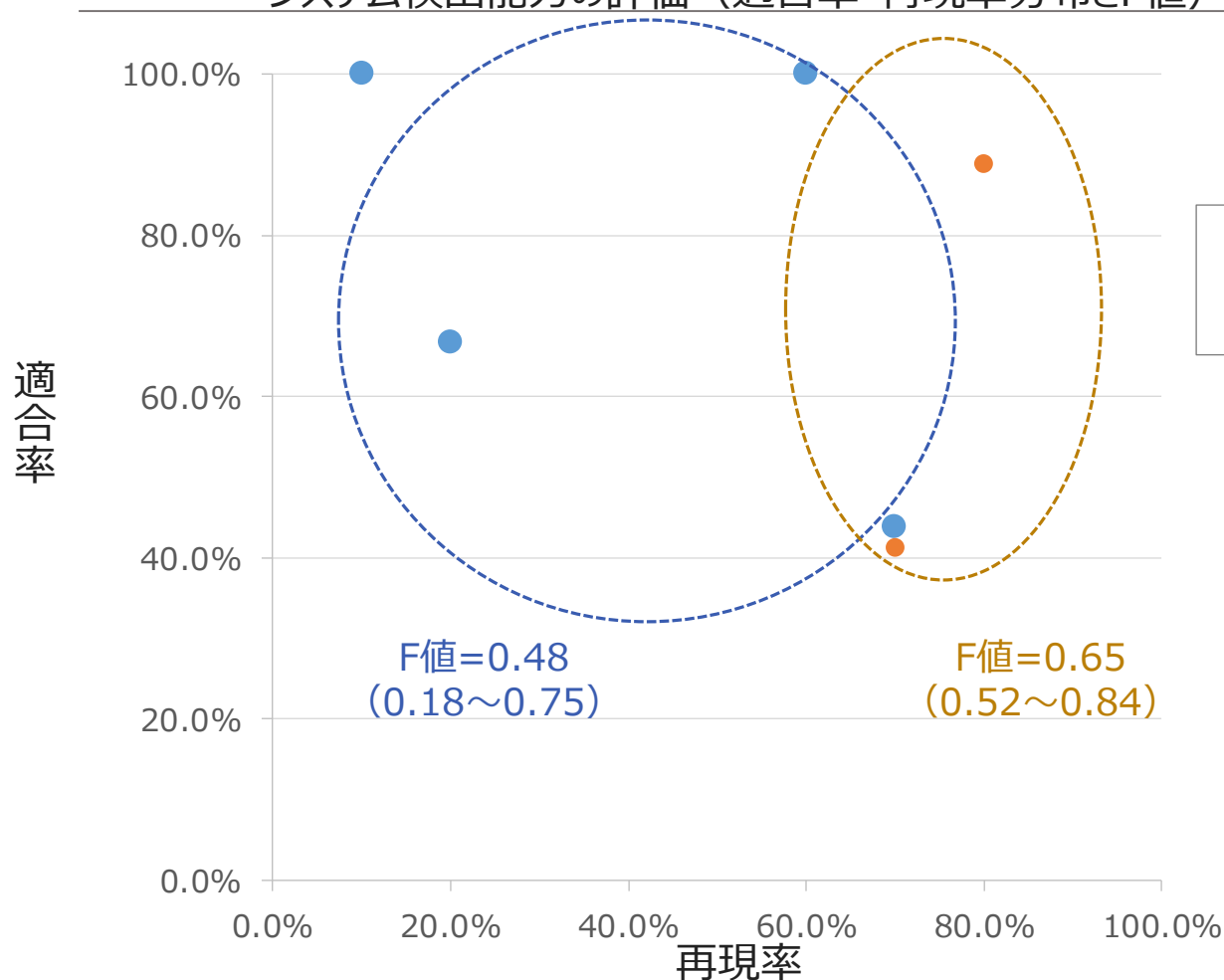
分散比 = 新方法【AI】の分散 ÷ 現行方法の分散

4.2 結果②お客様との通話記録 【高度化】 適合率・再現率（システム検出能力）

証券様

適合率と再現率の分布より、システム検出能力の増加が確認された。（F値で0.48⇒0.65へ増加）

システム検出能力の評価（適合率・再現率分布とF値）



$$F\text{値} = \frac{\text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$$

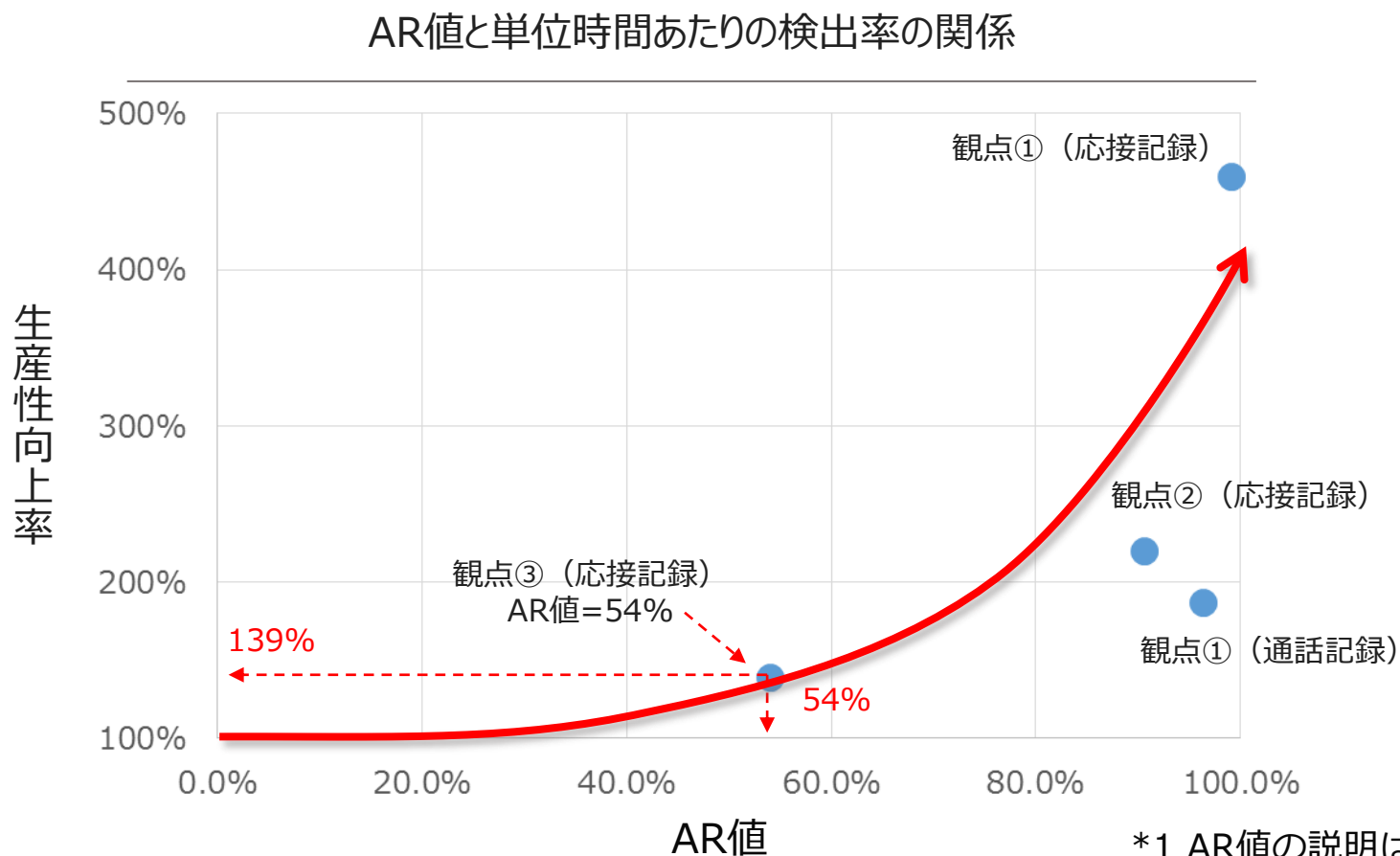
※右上にマップされるほど検出能力は高い

1. 本実証実験の概要
 - 1.1 本実証実験の概要および体制図
 - 1.2 実証実験ハブ参加検討に至った背景
 - 1.3 実証実験における目的や目指すところ
2. 実験テーマと実施方法
3. 実験条件
 - 3.1 実験条件詳細
 - 3.2 KIBITによる事前精度検証
4. 実験結果
 - 4.1 結果①応接記録
 - 4.2 結果②お客様との通話記録
5. 結論と考察
 - 5.1 結論と考察
 - 5.2 所感
 - 5.3 監督指針上、確認させていただきたいポイント
 - 5.4 実験参加者のアンケート回答内容

5.1 結論と考察

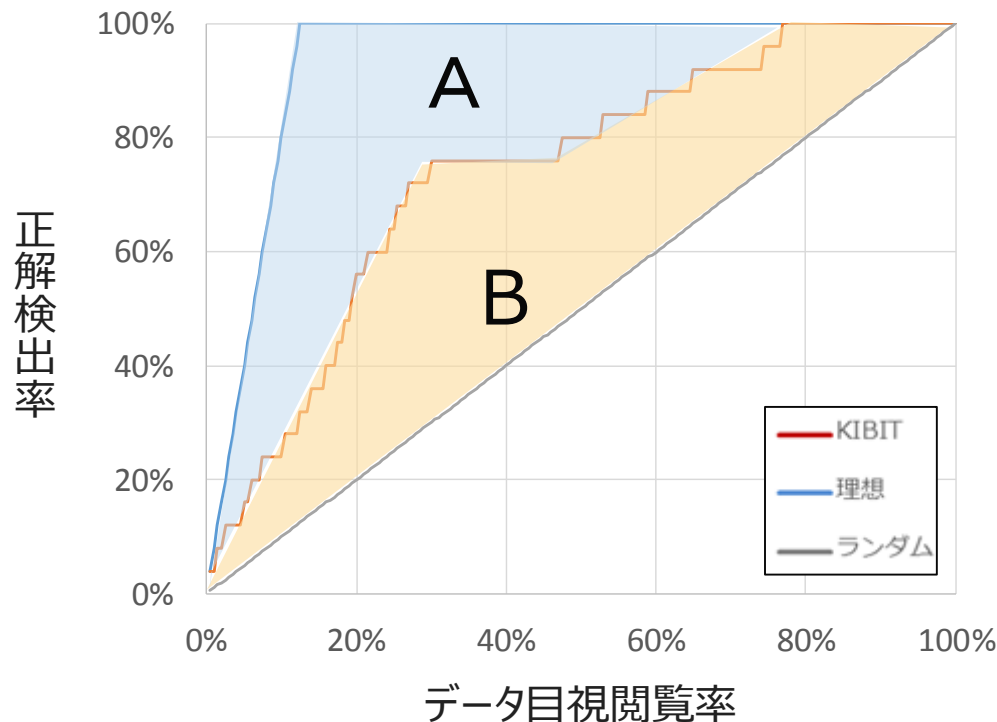
5.1 結論と考察 – モデル精度と生産性向上率

4社の生産性向上率と、使用した学習済みモデルの精度(AR値*1)を以下のグラフに示す。
AR値が高い程、生産性向上率は高まることが確認できる一方、AR値54%程度の学習済みモデルでも、人による現行方法と比べ139%と生産性の向上が図れることを確認できた。
また学習済みモデルを精度向上のためチューニングしていく事で、検出率を上げることが可能である。



AR値とは、以下のグラフから描かれる面積を使って計算される、AI計算モデルの理想モデルへの近さを表す指標である。

AR値の定義：閲覧率と正解検出率



$$\text{AR値} = \frac{B}{A+B}$$

※AR値が100%に近いほど、理想に近い

5.1 結論と考察 – AIを用いたチェックの精度と効率（1/2）

実験時間内にチェックが完了しなかったケースにおいては実験中と同等の精度・速度で完了までチェックしたと仮定。全件チェック完了時での推定正解検出率と比べ、新方法【AI】での検出率は銀行様・証券様共に同等以上であった。

■ 各観点での 被験者の一人あたり平均 実施時間・正解検出率		実験終了時		全件チェック完了時(推定)	
		新方法AI（実測値）		現行方法(推定値)	
		時間 (分)	正解検出率 (%)	時間 (分)	推定正解検出率 (%)
銀行	・応接記録	74分	65%	118分	61%
	・通話記録	60分	75%	131分	72%

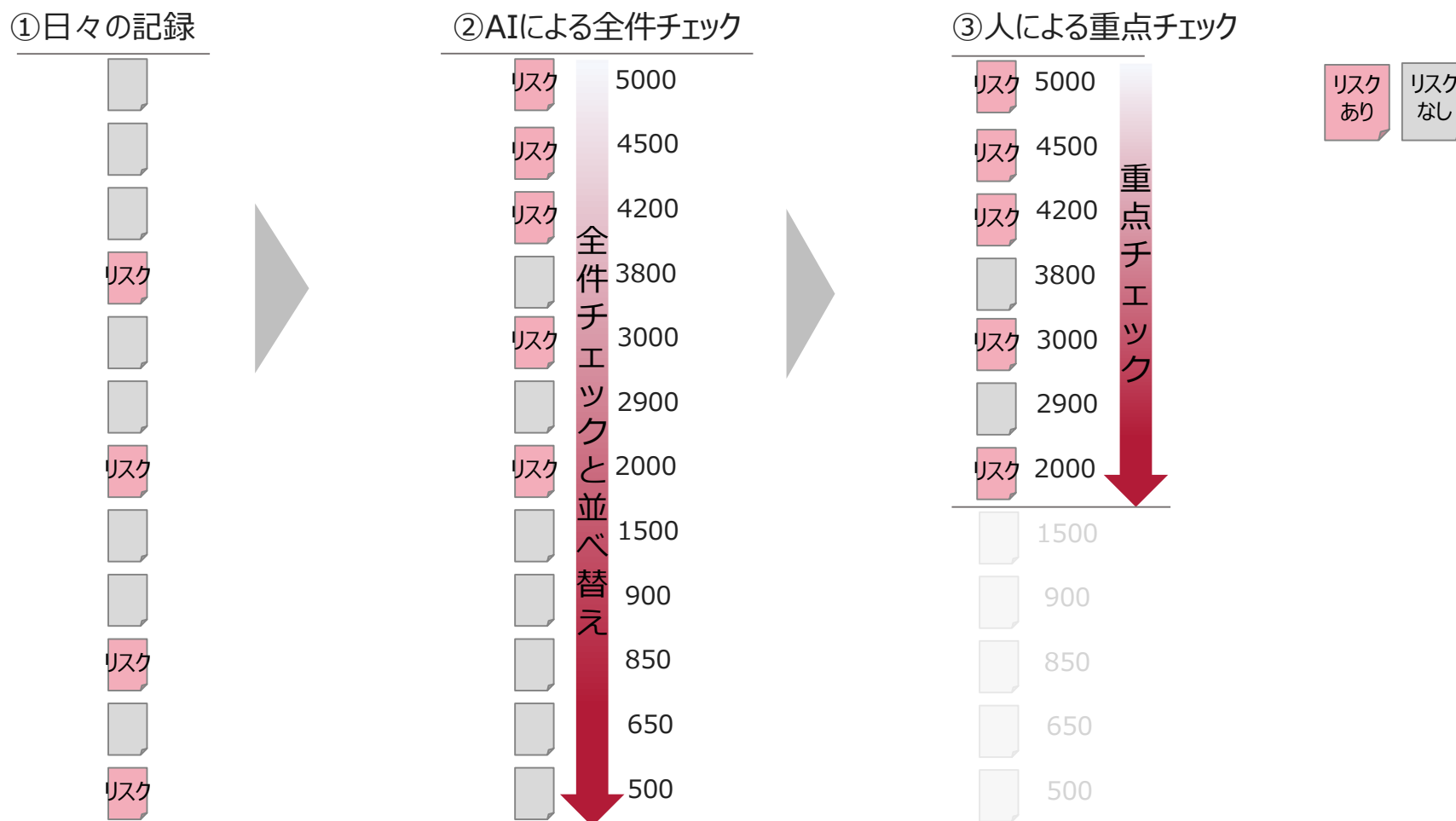
全件チェックと同等以上の正解検出率を維持しながら
チェック業務の時間短縮が見込める

※正解検出率=正解検出件数÷用意した正解数

5.1 結論と考察 – AIの効果的な使い方（2/2）

AIを導入してチェック業務効率を高めていくためには、**全データのチェック業務はAIに委ね、高リスクのデータに対して人の目によるチェックをする運用**が適している。

AIによる効果的と考えら得る業務フロー



	測定指標	結果（サマリ）
精度	正解検出率	・現行方法44%、新方法【AI】65%と、AIを使用することで正解検出率は高まった。
省力化・効率化（生産性）	時間当り正解検出件数	・現行方法（4.1件/時間）、新方法【AI】（9.1件/時間）とAIを使用した新業務は、 時間当りで5.0件検出能力が高かった。
	生産性向上率	・現行方法・新方法【AI】の比較で、 生産性向上率は225%と増加した。 ・スキル別では 低スキル参加者への効果が高い 傾向がみられた。
標準化（作業品質）	分散比	・検出率の増加により、全体ではばらつきの減少を確認できなかったものの、 観点別では分散比は縮小し、標準化の傾向を確認 することができた。
高度化（検出能力）	適合率・再現率分布（+F値）	・現行方法より、新方法【AI】の方が 検出能力が高い ことが確認された。 ※F値：現行（0.44）、新方法【AI】（0.66）、 増加率150%

結論と考察

- ・3観点全体で高い生産性向上の効果を確認され、特に業務経験がそれ程多くない実験者（経験普通）に対して、AIによる補助効果（検出件数増加と標準化作用）が強く現れていたことが確認できた。
- ・応接記録の確認業務を一連のシステムとした場合、現行システムとAIを使った新システムで高度化の効果（少ない数を見て、正解を見つける）があることが確認された。
- ・本実験では、応接記録に関わる3観点のみの検証を行った。今後、業務全体を高度化するために、他の観点のAI化およびAI化に合わせた制度・運用方法の検討が必要である。

	測定指標	結果（サマリ）
精度	正解検出率	・現行方法40%、新方法【AI】75%とAIを使用することで正解検出率は高まった。
省力化・効率化（生産性）	時間当り正解検出件数	・現行方法（4.0件/時間）、新方法【AI】（7.5件/時間）とAIを使用した新業務は、 時間当りで3.5件検出能力が高かった。
	生産性向上率	・現行方法、新方法【AI】の比較で、 生産性向上率は188%と増加した。 ・スキル別では 高スキル参加者への効果が高い 傾向がみられた。
標準化（作業品質）	分散比	・ 分散比は6%へと縮小し、標準化の傾向を確認 することができた。
高度化（検出能力）	適合率・再現率分布（+F値）	・現行方法より、新方法【AI】の方が 検出能力が高い ことが確認された。 ※F値：現行（0.48）、新方法【AI】（0.65）、 増加率135%

結論と考察

- ・本実験では、有効集計対象データが、現行方法4名分、新方法2名分と、数が少ないことに留意する必要があるものの、高い生産性向上の効果を確認することができた。
- ・通話記録の確認業務を一連のシステムとした場合、現行システムとAIを使った新システムで高度化の効果（少ない数を見て、正解を見つける）があることが確認された。
- ・今後、通録チェック業務を高度化するに当たり、テキスト判断して良いもの、音声判断して良いものは何か等の、AI化に合わせた制度・運用方法の検討が必要である。
- ・実運用において、自社での1日約10万件の通話録音データの確認業務に適用した場合、人に比べて600倍程度の効率化が図れる想定である。

5.1 結論と考察 – 実証実験まとめ

今回の銀行様、証券様における実証実験の結果を以下にまとめる。

実験結果		考察	人とKIBITの比較結果	
全件チェック	精度	KIBITの検出率は 人と同等以上 であった	KIBITは人と同等以上の検出判断ができる（現状のチェック業務の検出クオリティは維持できる）	同位 人工知能 KIBIT
	効率化	人に比べてKIBITの検出時間は 42% 短かった	KIBITは人に比べて短時間で検出業務を完了できる	優位 人工知能 KIBIT
（単位時間当り） サンプルチェック	精度	人に比べてKIBITの方が 単位時間当り正解検出数は 2倍 であった	KIBITは人に比べて高い検出数を実現できる	優位 人工知能 KIBIT
共通	標準化	業務経験の多寡による 検出能力のバラつきは、KIBITの方が少なかった。 また 検出能力もKIBITを用いた方が高かった	業務経験の多寡によるチェック品質のバラつきや人材育成、採用難という課題に対処できる	優位 人工知能 KIBIT



5.2 所感

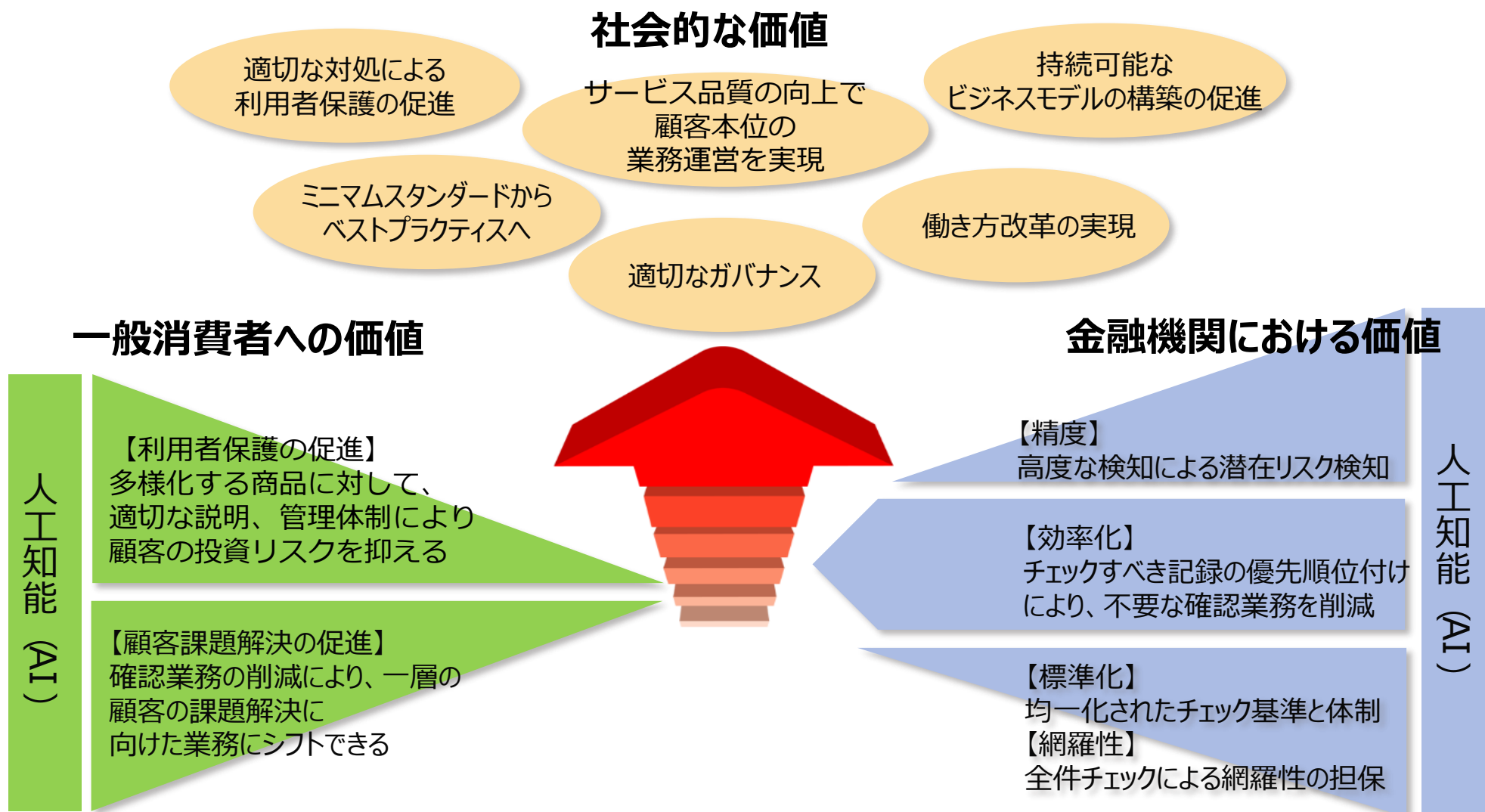
今回の実証実験の結果を踏まえ、以下総括を記す。

昨今の金融業界では一般消費者のニーズや金融商品の多様化のため、これまで以上に一般消費者への適切で、丁寧な商品説明やお客様対応が求められている。また今後、より一層お客様に寄り添った顧客本位の業務運営が求められるが、働き方改革を意識した業務運営も考慮する必要性に迫られている。

今回の実証実験ハブでは、こういった金融機関における業務の高度化、効率化といった課題を人工知能で解決できるかを検証した。実証実験の結果としては、注意喚起が必要な取引や活動の記録、またお客様からのお申し出の記録をこれまでの人だけによる確認と比べ、人工知能は人が持つ暗黙知を十分学習し、同等の検知ができることを証明できた。また人工知能は全ての記録に対して網羅的に高速で内容をチェックし、リスクの高い記録を検知する事を証明できた。

金融機関の実業務での人工知能の利用を想定すると、他観点でのモデル構築・精度検証等をおこなうことで、これまで以上にリスクの高い記録を網羅的に確認する事ができ、業務の高度化が期待できると考える。また不要な確認業務を削減する事で、業務の効率化も期待できると考える。

チェック業務で人工知能を用いる事で、一般消費者と金融機関における価値から生み出される社会的な価値。



5.3 監督指針上、確認させていただきたいポイント

人工知能をチェック業務で利用する際の懸念点と金融庁様の見解

1. チェック業務にてAIによる一次チェック、人による二次チェックを行う運用は問題があるか。
(金融商品販売時のコンプライアンス関連のチェック業務においてAIを補助的に活用することは有効であると考えており、AIによる一次的チェックを行い、リスクの高い取引を絞り込んだうえで、検証の担い手が二次的に絞り込んだ分のみのチェックを行う運用をすることについて問題はあるか。)
2. 各金融機関でAIによる判定基準を独自に設定し運用する事は問題があるか。
(スクリーニングツールという前提にて使用をする。その際の「●点以下は見ない、上位●%のみを見る」といった基準は、適量の過去事例におけるスコアリング結果やバックテスト等に基づき各金融機関の独自基準を持って設定することで問題はないか。)
3. 各金融機関でAIの学習済みモデルの信頼性の確認の周期を独自に設定する事は問題があるか。
(作成モデルの精度については、一定周期毎に全量チェックをすることで検証を行うと同時に、その信頼性についての維持につとめるものとする。その周期については上記の基準と同様、各金融機関毎の基準に沿って設定するものとしたいが問題はないか。)



■ 金融庁様見解

例えば、AIによる判定基準や学習済みモデルの信頼性等に関する検証を合理的な方法・間隔で行う等、適切な運用がなされているのであれば、法令・監督指針上、金融機関による確認業務に関し、AIによる一次確認を介する運用を行うことに特段の問題はないと考えられる。

5.4 実験参加者のアンケート回答内容

【アンケート1、2】 KIBITによるスコア付けにより、モニタリングの効率は向上したと感じますか。

【結果】

主な回答理由

評 価	人 数
5 (とても向上した)	1
4 (やや向上した)	12
3 (どちらとも言えない)	7
2 (あまり向上しなかった)	1
1 (全く向上しなかった)	0

- AIによるスコアリングの高い方に、問題のある応接記録が集中していたと感じました。

- KIBITによるスコア付けがあることにより、同じ時間でも注意深く見ながらも回答数が増加し、見落としも少ない結果となり、生産性向上に寄与度が上がっていると感じました。

- スコア順に並んでいるので、優先順位を付ける作業が軽減されると思います。一定基準のスコア以上を確認する等にできれば、モニタリング作業そのものの効率化に繋がると思います。

- 点数に引きずられる部分があったと感じました。

【アンケート3、4】 現行の方法と比べて、KIBITスコア付の方がモニタリングはやり易いですか。

【結果】

評 価	人 数
は い	16
いいえ	5

主な回答理由

- スコアが高い事例は「必要な要素」が多く含まれているであろうとの一応の安心感があり、何もなくモニタリングする場合よりは安心度がある分、やり易いと感じました。
- スコア順に並んでいるため、優先順位が高い案件から着手でき、作業指示も出し易くなると思います。
- スコアがあるとより慎重にモニタリングする意識が高まり、問題点が把握し易くなると感じました。
- 現行のモニタリングでは、一つの観点のみでのモニタリングではないので、一概に比較できないと感じました。
- KIBITのスコア付の有無に関わらず、テキストのモニタリングのみでは内容を詳細に把握するのは困難だと思います。



この資料は下記により作成されました。

株式会社FRONTEO

108-0075 東京都港区港南2-12-23 明産高浜ビル

TEL : 03-5463-6375 FAX : 03-6825-3881

東証マザーズ上場 | Nasdaq (ナスダック) 上場

本資料は、株式会社FRONTEOの承諾なしに貴社以外の第三者の閲覧に供することは出来ません。

本資料のいかなる部分も、一切の権利は株式会社FRONTEOに帰属しており、電子的または機械的な方法を問わず、いかなる目的であれ、無断で複製・転送されることは固くお断りいたします。