

SVMによる整形外科領域の慢性疼痛原因診断に関する性能評価

Experimental Evaluation of SVM to Diagnose Causes of Chronic Pain in Orthopedics

○岩崎翔太¹, 平石秀史²

*Shota Iwazaki¹, Hidefumi Hiraishi²

Abstract: This study explores the application of Support Vector Machines for chronic pain diagnosis toward Explainable AI in this domain. We utilized weighted items of PainDETECT and BS-POP questionnaires as feature vectors, with item weights optimized via hill climbing. Our proposed framework achieved an accuracy of more than 70%.

1 はじめに

整形外科での典型的な主訴の一つに疼痛がある。疼痛は、世界の人口の20%以上が何らかの疼痛を感じているとの報告がある程、普遍的な症状である[pain information]。疼痛は、発生する原因によって、侵害受容性疼痛、神経障害性疼痛、原因不明に分類される。疼痛診断では、PainDETECT[1]やBS-POP[2]という問診票の回答を診断材料に用いる場合がある。本研究では、これらの問診票を用いたAIによる診断支援の可能性について検討する[3]。

医療診断をAIを用いて行う際、診断の根拠を明示するために説明可能なAI(XAI)を構築することが求められる。また、領域によっては医療機関が保有し利用が可能なデータ量が限定的な場合も多い。本研究では、限定的なデータ量で一定の精度が達成でき、また結果の解釈も比較的容易なサポートベクトルマシン(SVM)に着目し実験を行う。

1.1 サポートベクトルマシン(SVM)

SVMは、与えられた学習データに対し、最適な超平面を定めて分離する機械学習モデルである。超平面を定める際に、境界と最も近い各クラスのデータのみを重視し、このデータとの距離が最大になるように定めることで、少ないデータセットでも高い汎化性能を達成できる可能性をもつモデルである[4]。

SVMには線形分離可能なデータに用いられるハードマージンSVMと、誤りを許容することで線形分離可能でないデータにも適用可能なソフトマージンSVMがある。ハードマージンSVMは学習データ数 n 、特徴量ベクトル \mathbf{x}_i ($i = 1, 2, \dots, n$)、各データのラベル t_i ($i = 1, 2, \dots, n$)を用いて、次の最適化問題として定式化される。

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{subject to} \quad t_i (\mathbf{w}^\top \mathbf{x}_i + b) \geq 1 - \xi_i \quad (i = 1, \dots, n)$$

ソフトマージンSVMはコストパラメータ C 、誤分類の程度を表現するパラメータ ξ_i ($i = 1, 2, \dots, n$)を導入し、次

のように定式化される。

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i$$

subject to $t_i (\mathbf{w}^\top \mathbf{x}_i + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (i = 1, \dots, n)$

1.2 カーネル法

SVMの線形分離可能でない場合には、カーネル法により入力ベクトル \mathbf{x} を高次特徴空間 $\phi(\mathbf{x})$ に写像し、その空間で超平面による分離を行う。写像を明示せずに内積のみをカーネル関数 $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^\top \phi(\mathbf{x}_j)$ として計算することで、計算量を抑えつつ非線形境界を表現する。代表的なカーネルは以下の4つである。

- (1) 線形 (Linear) : $k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^\top \mathbf{x}_j$
- (2) 多項式 (Poly) : $k(\mathbf{x}_i, \mathbf{x}_j) = \gamma(\mathbf{x}_i^\top \mathbf{x}_j + c)^d$
- (3) ガウス (RBF) : $k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$
- (4) シグモイド (Sigmoid) : $k(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^\top \mathbf{x}_j + c)$

以上のカーネルについて、パラメータの調整を行い、性能比較や説明可能性を統合評価して最適なカーネルを選択する。

2 実験詳細

本研究では、PainDETECT(質問項目:13)、BS-POP(質問項目:18)、及び両者を結合したFUSION(質問項目:31)を用い、侵害受容性疼痛、神経障害性疼痛、原因不明の3値を目的変数とした分類モデルを構築した。生データのデータ数は801で、クラス内訳(侵害受容性疼痛/神経障害性疼痛/原因不明)は61%/32%/7%であった。データクレンジングでは、明らかな入力誤りや欠損を点検した上で、解析バイアスの増幅を避ける目的で欠損を含むデータは削除し(データ数:540, クラス内訳61%/33%/6%)、その後、PainDETECT, BS-POP共に質問項目のスコアレンジ

が異なるため、特徴量を平均 0・分散 1 に標準化した。そして、ソフトマージン SVM を用い、線形カーネル、多項式カーネル、ガウスカーネル、シグモイドカーネルの判定精度を比較した。各特徴量が予測に与える影響の強さは同じではないと考え、各設問へ重み付け ($x'_j = \alpha_j x_j$) を導入した。重み付けの設定を自動化するため、 α を山登り法により調整を行った (初期値: 1, ステップ幅: 0.01, 反復回数: 1000)。さらに、各カーネルのハイパーパラメータ (C, γ, d, c) はグリッドサーチで探索した。汎化性能は 5 分割交差検証で評価した。

- (1) 標準 SVM(SVM+kernel)
- (2) 山登り法による重み最適化 SVM(SVM+w+kernel)

の 2 つの条件を、正答率、感度、特異度の評価指標のもと、精度を比較するものとする。

3 実験結果

実験結果を表 1 (a) (b) (c) に示す。表には山登り法により重みを設定した結果のうち、最も正答率を高かったものを記載する。

PainDETECT は各質問項目のスコア総和により疼痛原因を推測するが、最終的な医師の診断と異なりうる。今回用いたデータでは、スコア総和による判定の正答率は 55.4%、感度/特異度は侵害受容性疼痛 80.1%/38.8%、神経障害性疼痛 15.3%/94.2%、原因不明 21.9%/81.9%であった。これに対し、機械学習モデルとして SVM を用いることで、PainDETECT のスコア総和による判定より高精度になることがわかる。特に、神経障害性疼痛の感度が約 15%から約 50%近く向上させることができた。BS-POP は整形外科患者における精神医学的問題を知ることにかけている質問表であるため、神経障害性疼痛の診断制度は高いと予測していたが、PainDETECT を下回った。

4 考察と今後の展望

本研究では、標準 SVM に対して質問項目別の重み付けを導入し山登り法で最適化することで、小幅ながら一貫した精度向上が確認できた。これは各質問項目の情報量が均一ではないことを示唆し、特徴量重み最適化の有効性を裏付ける結果である。一方でクラス別の結果を見ると、

侵害受容性疼痛の感度は高いが、神経障害性疼痛の感度が相対的に低いことから、これが正答率に影響している可能性がある。

今後は、単に全体正答率を最大化する目的関数で山登りを行うのではなく、神経障害性疼痛に焦点を当てた目的関数で最適化を行いたいと考えている。また、進化計算アルゴリズムを用いた重み付けの探索や主成分分析 (PCA) で冗長な成分を把握し重要度の高い成分に重みを集中させる設計を検討する。さらに、ハイパーパラメータはより細域でのチューニングにより信頼性を高める。これらを通じて、特に神経障害性疼痛の検出性能を向上させ、全体として正答率 80%以上の達成を目標とする。

参考文献

- [1] Rainer Freynhagen et al. “Pain DETECT: a new screening questionnaire to identify neuropathic components in patients with back pain”. In: Current medical research and opinion 22.10 (2006), pp. 1911–1920.
- [2] Katsuhiko Sato et al. “Liaison consultation psychiatry for spine and spinal cord disease; brief scale for evaluation of psychiatric problems in orthopaedic patients (BS-POP)”. In: Clin. Orthop. Surg 35 (2000), pp. 843–852.
- [3] Satoshi Suzuki et al. “Pain Screening in Orthopedics Using Machine Learning”. In: TENCON 2024-2024 IEEE Region 10 Conference (TENCON). IEEE. 2024, pp. 624–627.
- [4] 田村孝廣. やさしく学べるサポートベクトルマシン: 数学の基礎と Python による実践. オーム社, 2022. URL: <https://ndlsearch.ndl.go.jp/books/R100000002-I032497308>.

		感度/特異度 (%)						感度/特異度 (%)						感度/特異度 (%)			
方式	正答率 (%)	侵害受容性	神経障害性	不明	方式	正答率 (%)	侵害受容性	神経障害性	不明	方式	正答率 (%)	侵害受容性	神経障害性	不明			
SVM-Linear	71.3	90.9/61.9	65.7/87.7	0/100	SVM-Linear	63.5	70.0/16.7	20.0/97.3	0/100	SVM-Linear	71.3	80.3/52.4	54.3/78.1	0/100			
SVM-Sigmoid	72.2	83.3/40.5	48.6/84.9	0/100	SVM-Sigmoid	62.0	93.9/28.6	28.6/91.8	0/100	SVM-Sigmoid	70.6	90.9/47.6	54.3/90.4	0/100			
SVM-RBF	71.9	90.9/45.2	51.4/90.4	0/100	SVM-RBF	63.5	86.4/26.2	25.7/84.9	0/100	SVM-RBF	71.9	92.4/47.6	54.3/91.8	0/100			
SVM-Poly	72.6	89.4/52.4	60.0/89.0	0/100	SVM-Poly	63.3	97.0/14.3	17.1/97.3	0/100	SVM-Poly	71.7	95.5/42.9	51.4/95.9	0/100			
SVM+w-Linear	74.1	89.4/61.9	69.4/88.9	0/100	SVM+w-Linear	65.6	90.9/28.6	34.3/91.8	0/100	SVM+w-Linear	74.3	87.9/66.7	74.3/86.3	0/100			
SVM+w-Sigmoid	72.0	87.9/61.9	69.4/87.5	0/100	SVM+w-Sigmoid	63.9	93.9/26.2	25.7/91.8	0/100	SVM+w-Sigmoid	71.7	92.4/54.76	62.9/91.8	0/100			
SVM+w-RBF	71.5	89.4/66.7	74.3/87.7	0/100	SVM+w-RBF	63.2	92.4/33.3	34.3/90.4	0/100	SVM+w-RBF	73.3	93.4/61.9	68.6/91.8	0/100			
SVM+w-Poly	71.9	89.4/61.9	69.4/88.9	0/100	SVM+w-Poly	63.7	93.9/30.2	34.3/93.2	0/100	SVM+w-Poly	73.3	87.9/64.3	69.4/86.1	0/100			

(a) PainDETECT (b) BS-POP (c) FUSION
表 1: 実験結果