

M34a 太陽フレア予測のための特徴量の重要度と評価手法考察

西塚直人, 久保勇樹, 田光江, 亘慎一, 石井守 (情報通信研究機構 電磁波研究所), 杉浦孔明 (情報通信研究機構 先進的音声翻訳研究開発推進センター)

太陽フレア発生機構は、太陽物理の長年の未解決問題である。我々は、活動領域の光球ベクトル磁場と彩層発光、および軟X線放射量をもとに特徴量を抽出し、統計的機械学習によってフレア発生を予測するモデルを開発した(日本天文学会 2016年春季年会 西塚ら)。観測データベースは2010年から2015年に取得されたSDO衛星の光球ベクトル磁場、視線方向磁場、彩層低部発光1600Å、およびGOES軟X線放射のデータをもとにし、活動領域を自動検出し、さらに特徴量データベースを作成する。50程の特徴量をもとに機械学習してフレア予測を行い、人手による予測よりも精度の良い予報を達成することに成功した。

本講演では、何がフレア予報にとって重要な特徴量であるかに着目する。機械学習法の1つ、Extra Random Treesにより特徴量の重要度を測ることができる。その結果、おおよそフレア履歴、前日のX線量、磁気中性線の情報、磁束量、彩層発光面積、磁場強度という順で重要であることがわかった。さらにリアルタイム予報の評価法について3通りの試行実験を行った。1つ目は全データをフルシャッフルして学習・テスト用データへ分割。2つ目は1週間のうち決まった曜日1日をテスト用データ、他の6日分を学習用データとした分割。3つ目は1週間のうち決まった曜日1日をテスト用として、その前後の曜日を除いた4日を学習用データとした。3つ目のサンプリングでは、完全に学習用データのフレアとテスト用データのフレアとは別物になり、予測が最も難しい。フルシャッフルではTSSが0.9を超える精度であるのに対し、2つ目の評価法では0.5-0.9、3つ目では0.2-0.3と下がる。本講演では、特徴量の重要度と、評価手法についての考察結果を述べる。