

エルニーニョ現象・ラニーニャ現象予測システム

UECスクールAチーム
川瀬冬馬 鈴木凜央
笠巻 雫 須藤実里

○概要

エルニーニョ現象・ラニーニャ現象は異常気象の要因となる可能性があり、これらがいつ発生するかを知ることは防災の観点から重要ですが、長期予測は現在困難です。そこで過去の海面水温のデータ及びエルニーニョ現象・ラニーニャ現象と相関関係にあるデータを用いて、長期的にこれらの現象を予測するシステムを作成しました。

○研究内容

エルニーニョ・南方振動は潮汐8.6年運動と連動しています。(右下図潮汐年とエルニーニョ現象・ラニーニャ現象が起きた年の対応表) また、黒点とこれらの現象には弱い相関があることを発見しました。これらのことより気象庁・国立天文台の過去54年分の海面水温・潮汐・海面水位・黒点相対数のデータを基に、エルニーニョ現象・ラニーニャ現象の基となる海水温をLSTMを用いて長期予測するシステムを開発しました。

海水温・潮汐・海面水位・黒点相対数のデータから用意したCSVファイルを加工しました。pytorchを用いて作成したLSTMのモデル(以下model1とする)にこれらのデータを入力し、学習させました。学習したモデルでデータを用い、海水温を予測しました。

Tendency	0	1	2	3	10	12	13	14
1690	1710	1712	1719	1721	1722	1728		
1727	1728	1730	1737	1738	1740	1743		
1748	1757	1768	1764	1769	1769	1769		
1784	1785	1787	1774	1776	1777	1780		
1783	1784	1786	1793	1795	1796	1799		
1802	1803	1806	1819	1814	1815	1818		
1820	1821	1823	1830	1832	1833	1836		
1839	1840	1842	1849	1851	1852	1855		
1858	1859	1861	1868	1870	1871	1874		
1876	1877	1879	1886	1889	1889	1892		
1895	1896	1898	1905	1907	1908	1911		
1913	1914	1916	1923	1925	1926	1929		
1932	1933	1935	1942	1944	1945	1948		
1951	1952	1954	1961	1963	1964	1967		
1969	1970	1972	1979	1981	1982	1985		
1988	1989	1991	1998	2000	2001	2004		
2006	2007	2009	2016	2018	2019	2022		
2025	2026	2028	2035	2037	2038	2041		
2044	2045	2047	2054	2056	2057	2060		

加えて、エルニーニョ現象と関わりの深い気象庁のエルニーニョ監視海域(NINO.3)、西太平洋熱帯域(NINO.WEST)、インド洋熱帯域(IOBW)においての海面水温、基準値、基準値との差、五か月移動平均値、南方振動指数の過去74年分のデータ及びいつエルニーニョ現象・ラニーニャ現象が起きたかのラベルデータを用いました。(下図各海域) これらの特徴量データを年・月で統合してつのcsvファイルにします。ラベルデータも同様の形式にしてラニーニャ現象発生時を0、エルニーニョ現象発生時を1、通常時Neutralを2として判別します。

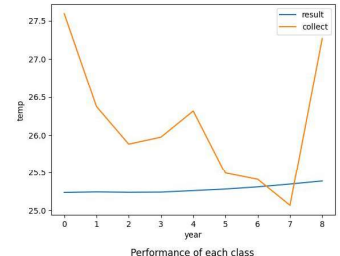


○今後の展望

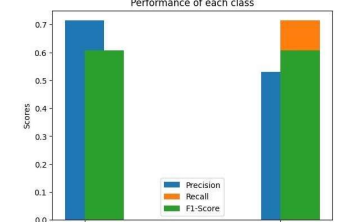
今回のシステムの試行結果から得られた課題を解決するためAIに対してさらに多くのデータを学習させることで、より正確に予測できるよう改善を図ります。また海面水温データを用いた予測に加えタヒチとダーウィンの気圧差を基に算出される南方振動指数(SOI)の予測を導入することなどで、エルニーニョおよびラニーニャ現象の発生時期をより正確に予測し、全体の予測精度の向上を目指していきます。以上の取り組みによって、これらの現象の予測がより正確になり、実際の気象予測に役立つシステムになることを期待しています。

○試行結果

Model1に過去54年分の海水温データを入力し、エルニーニョ現象・ラニーニャ現象の長期予測を行いました。試行では、各年の海水温データに基づいて複数回の予測を行い、モデルの精度を確認しました。結果として、未知のデータと予測結果の評価関数RMSEは5.07となっています。(右図予測結果と実際の数値のグラフ)



Model2に74年分のデータのうち直近20年をテスト用、残りの54年を学習用にして、LightGBMを用いてモデリングしました。精度向上に向けてオーバーサンプリングとクロスバリデーションを併用しました。モデリング時の精度は0.9913だったのに対し、テストデータでの精度は0.608でした。分類レポートでは再現率がクラス1(EINino)が高く、精度はクラス0(LaNina)が高いという特徴が出ました。



○考察

今回の試行結果を踏まえ、いくつかの重要な示唆が得られました。まず、model1においては、未知のデータと予測結果の評価関数RMSEは5.07で、正確に予測出来ていないということがわかります。同様にグラフでも、予測が出来ていないことがわかります。その対策として、データセットや学習のアルゴリズムの見直しが必要であると考えられます。これらの結果から、モデルの学習回数やデータセットの量について、さらなる調整が必要であることが示唆されます。今後の改善に向けて、これらの要因を考慮し、精度向上を目指した取り組みを進める予定です。一方、model2においては、テスト時も一定の精度を示しているものの精度が訓練時に比べて低くなっていることから過学習や訓練データとテストデータの偏りがある可能性が高いと考えられます。対策としては、オーバーサンプリングやデータバランスの調整をさらに進めることや、XGBoost等他モデルを追加することが挙げられます。

参考文献: <https://www.aori.u-tokyo.ac.jp/research/news/2018/20181012.html>
<https://disassemble-channel.com/1stm-multivariate-time-series/>
https://www.data.jma.go.jp/gmd/kaiyou/db/tide/sl_area/sl_rtrend.html
https://solarwww.mtk.nao.ac.jp/ip/db_sunspot.html
https://www.data.jma.go.jp/gmd/cpd/data/el_nino/index/nino3abs.html
https://www.data.jma.go.jp/cpd/data/el_nino/index/datab.html
https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php