Accuracy

Models	basque-DO-S_DO_V_AUX	basque-DO-S_IO_DO_V_AUX	basque-IO-IO_S_V_AUX	basque-IO_IO_S_DO_V_AUX	basque-S-IO_S_DO_V_AUX	basque-S-S_DO_V_AUX	basque-S-S_IO_DO_V_AUX	basque-S-S_V_AUX	hindi-S_ne_O_V	hindi-S_ne_PossPRN_O_V	hindi-S_PossPRN_PossN_O_V	hindi-S_O_V	hindi-S_PossPRN_O_V	hindi-S_PossPRN_PossN_O_V	swahili-N_of_Poss_D_A_V	swahili-N_of_Poss_D_AP_ni_AN	swahili-N_of_Poss_D_AP_V_ni_AN	swahili-N_of_Poss_D_ni_A	swahili-N_of_Poss_D_V	swahili-N_of_Poss_V
human	1.000 (0.996-1.000)	0.920 (0.901-0.936)	0.980 (0.969-0.988)	0.820 (0.795-0.843)	0.860 (0.837-0.881)	0.700 (0.671-0.728)	0.920 (0.901-0.936)	0.860 (0.837-0.881)	0.840 (0.816-0.862)	0.920 (0.901-0.936)	0.820 (0.795-0.843)	0.820 (0.795-0.843)	0.880 (0.858-0.899)	0.740 (0.712-0.767)	0.740 (0.712-0.767)	0.780 (0.753-0.805)	0.640 (0.609-0.670)	0.660 (0.630-0.689)	0.780 (0.753-0.805)	0.860 (0.837-0.881)
mGPT-1.3B	0.996 (0.990-0.999)	0.978 (0.967-0.986)	0.910 (0.891-0.927)	0.973 (0.961-0.982)	0.991 (0.983-0.996)	0.897 (0.876-0.915)	0.750 (0.722-0.777)	0.913 (0.894-0.930)	0.956 (0.941-0.968)	0.966 (0.953-0.976)	0.974 (0.962-0.983)	0.831 (0.806-0.854)	0.914 (0.895-0.931)	0.933 (0.916-0.948)	0.408 (0.377-0.439)	0.564 (0.533-0.595)	0.527 (0.496-0.558)	0.594 (0.563-0.625)	0.444 (0.413-0.475)	0.668 (0.638-0.697)
mGPT-13B	0.991 (0.983-0.996)	0.988 (0.979-0.994)	0.794 (0.768-0.819)	0.994 (0.987-0.998)	0.999 (0.994-1.000)	0.861 (0.838-0.882)	0.916 (0.897-0.932)	0.836 (0.812-0.858)	0.950 (0.935-0.963)	0.929 (0.911-0.944)	0.944 (0.928-0.957)	0.805 (0.779-0.829)	0.922 (0.904-0.938)	0.938 (0.921-0.952)	0.481 (0.450-0.512)	0.594 (0.563-0.625)	0.542 (0.511-0.573)	0.610 (0.579-0.640)	0.494 (0.463-0.525)	0.709 (0.680-0.737)
bloom-560m	0.868 (0.845-0.888)	0.145 (0.124-0.168)	0.123 (0.103-0.145)	0.531 (0.500-0.562)	0.323 (0.294-0.353)	0.687 (0.657-0.716)	0.519 (0.488-0.550)	0.744 (0.716-0.771)	0.875 (0.853-0.895)	0.783 (0.756-0.808)	0.868 (0.845-0.888)	0.841 (0.817-0.863)	0.855 (0.832-0.876)	0.886 (0.865-0.905)	0.444 (0.413-0.475)	0.482 (0.451-0.513)	0.457 (0.426-0.488)	0.502 (0.471-0.533)	0.378 (0.348-0.409)	0.487 (0.456-0.518)
bloom-1b1	0.945 (0.929-0.958)	0.534 (0.503-0.565)	0.093 (0.076-0.113)	0.794 (0.768-0.819)	0.168 (0.145-0.193)	0.692 (0.662-0.721)	0.669 (0.639-0.698)	0.827 (0.802-0.850)	0.882 (0.860-0.901)	0.916 (0.897-0.932)	0.918 (0.899-0.934)	0.818 (0.793-0.841)	0.867 (0.844-0.887)	0.897 (0.876-0.915)	0.410 (0.379-0.441)	0.514 (0.483-0.545)	0.502 (0.471-0.533)	0.527 (0.496-0.558)	0.341 (0.312-0.371)	0.444 (0.413-0.475)
bloom-1b7	0.918 (0.899-0.934)	0.644 (0.613-0.674)	0.326 (0.297-0.356)	0.773 (0.746-0.799)	0.656 (0.626-0.685)	0.688 (0.658-0.717)	0.643 (0.612-0.673)	0.712 (0.683-0.740)	0.902 (0.882-0.920)	0.903 (0.883-0.921)	0.924 (0.906-0.940)	0.859 (0.836-0.880)	0.889 (0.868-0.908)	0.898 (0.878-0.916)	0.395 (0.365-0.426)	0.452 (0.421-0.483)	0.468 (0.437-0.499)	0.498 (0.467-0.529)	0.345 (0.316-0.375)	0.494 (0.463-0.525)
bloom-3b	0.981 (0.970-0.989)	0.754 (0.726-0.780)	0.452 (0.421-0.483)	0.836 (0.812-0.858)	0.753 (0.725-0.779)	0.853 (0.830-0.874)	0.773 (0.746-0.799)	0.882 (0.860-0.901)	0.858 (0.835-0.879)	0.913 (0.894-0.930)	0.917 (0.898-0.933)	0.829 (0.804-0.852)	0.896 (0.875-0.914)	0.904 (0.884-0.922)	0.378 (0.348-0.409)	0.481 (0.450-0.512)	0.463 (0.432-0.494)	0.539 (0.508-0.570)	0.339 (0.310-0.369)	0.494 (0.463-0.525)
bloom-7b1	0.981 (0.970-0.989)	0.920 (0.901-0.936)	0.724 (0.695-0.752)	0.932 (0.915-0.947)	0.928 (0.910-0.943)	0.869 (0.846-0.889)	0.864 (0.841-0.885)	0.801 (0.775-0.825)	0.869 (0.846-0.889)	0.904 (0.884-0.922)	0.921 (0.903-0.937)	0.839 (0.815-0.861)	0.901 (0.881-0.919)	0.887 (0.866-0.906)	0.400 (0.369-0.431)	0.501 (0.470-0.532)	0.471 (0.440-0.502)	0.511 (0.480-0.542)	0.405 (0.374-0.436)	0.510 (0.479-0.541)
bloom	0.997 (0.991-0.999)	0.983 (0.973-0.990)	0.914 (0.895-0.931)	0.993 (0.986-0.997)	0.995 (0.988-0.998)	0.923 (0.905-0.939)	0.908 (0.888-0.925)	0.922 (0.904-0.938)	0.905 (0.885-0.922)	0.942 (0.926-0.956)	0.942 (0.926-0.956)	0.908 (0.888-0.925)	0.941 (0.925-0.955)	0.919 (0.900-0.935)	0.416 (0.385-0.447)	0.523 (0.492-0.554)	0.484 (0.453-0.515)	0.548 (0.517-0.579)	0.432 (0.401-0.463)	0.552 (0.521-0.583)
xglm-564M	0.932 (0.915-0.947)	0.822 (0.797-0.845)	0.799 (0.773-0.823)	0.757 (0.729-0.783)	0.880 (0.858-0.899)	0.804 (0.778-0.828)	0.506 (0.475-0.537)	0.872 (0.850-0.892)	0.700 (0.671-0.728)	0.781 (0.754-0.806)	0.829 (0.804-0.852)	0.893 (0.872-0.911)	0.930 (0.912-0.945)	0.933 (0.916-0.948)	0.351 (0.321-0.381)	0.469 (0.438-0.500)	0.455 (0.424-0.486)	0.480 (0.449-0.511)	0.364 (0.334-0.395)	0.604 (0.573-0.634)
xglm-1.7B	0.955 (0.940-0.967)	0.917 (0.898-0.933)	0.905 (0.885-0.922)	0.903 (0.883-0.921)	0.967 (0.954-0.977)	0.884 (0.863-0.903)	0.654 (0.624-0.683)	0.837 (0.813-0.859)	0.829 (0.804-0.852)	0.852 (0.828-0.873)	0.877 (0.855-0.897)	0.943 (0.927-0.957)	0.961 (0.947-0.972)	0.958 (0.944-0.970)	0.461 (0.430-0.492)	0.531 (0.500-0.562)	0.514 (0.483-0.545)	0.558 (0.527-0.589)	0.547 (0.516-0.578)	0.689 (0.659-0.718)
xglm-2.9B	0.977 (0.966-0.985)	0.936 (0.919-0.950)	0.945 (0.929-0.958)	0.858 (0.835-0.879)	0.980 (0.969-0.988)	0.912 (0.893-0.929)	0.823 (0.798-0.846)	0.827 (0.802-0.850)	0.808 (0.782-0.832)	0.848 (0.824-0.870)	0.863 (0.840-0.884)	0.933 (0.916-0.948)	0.962 (0.948-0.973)	0.957 (0.943-0.969)	0.469 (0.438-0.500)	0.566 (0.535-0.597)	0.518 (0.487-0.549)	0.589 (0.558-0.620)	0.522 (0.491-0.553)	0.722 (0.693-0.750)
xglm-4.5B	0.607 (0.576-0.637)	0.548 (0.517-0.579)	0.413 (0.382-0.444)	0.285 (0.257-0.314)	0.550 (0.519-0.581)	0.473 (0.442-0.504)	0.517 (0.486-0.548)	0.582 (0.551-0.613)	0.742 (0.714-0.769)	0.786 (0.759-0.811)	0.811 (0.785-0.835)	0.881 (0.859-0.900)	0.926 (0.908-0.941)	0.946 (0.930-0.959)	0.387 (0.357-0.418)	0.502 (0.471-0.533)	0.477 (0.446-0.508)	0.525 (0.494-0.556)	0.412 (0.381-0.443)	0.601 (0.570-0.632)
xglm-7.5B	0.966 (0.953-0.976)	0.928 (0.910-0.943)	0.931 (0.913-0.946)	0.932 (0.915-0.947)	0.998 (0.993-1.000)	0.885 (0.864-0.904)	0.796 (0.770-0.821)	0.870 (0.848-0.890)	0.826 (0.801-0.849)	0.828 (0.803-0.851)	0.827 (0.802-0.850)	0.940 (0.923-0.954)	0.956 (0.941-0.968)	0.963 (0.949-0.974)	0.527 (0.496-0.558)	0.586 (0.555-0.617)	0.529 (0.498-0.560)	0.595 (0.564-0.626)	0.607 (0.576-0.637)	0.764 (0.736-0.790)
mbert	0.754 (0.726-0.780)	0.662 (0.632-0.691)	0.358 (0.328-0.389)	0.282 (0.254-0.311)	0.497 (0.466-0.528)	0.558 (0.527-0.589)	0.566 (0.535-0.597)	0.766 (0.738-0.792)	0.361 (0.331-0.392)	0.397 (0.367-0.428)	0.429 (0.398-0.460)	0.811 (0.785-0.835)	0.807 (0.781-0.831)	0.832 (0.807-0.855)	0.458 (0.427-0.489)	0.554 (0.523-0.585)	0.538 (0.507-0.569)	0.561 (0.530-0.592)	0.451 (0.420-0.482)	0.532 (0.501-0.563)
xlmr-base	0.661 (0.631-0.690)	0.596 (0.565-0.627)	0.597 (0.566-0.628)	0.465 (0.434-0.496)	0.640 (0.609-0.670)	0.659 (0.629-0.688)	0.384 (0.354-0.415)	0.728 (0.699-0.755)	0.764 (0.736-0.790)	0.831 (0.806-0.854)	0.836 (0.812-0.858)	0.864 (0.841-0.885)	0.854 (0.831-0.875)	0.874 (0.852-0.894)	0.488 (0.457-0.519)	0.518 (0.487-0.549)	0.504 (0.473-0.535)	0.513 (0.482-0.544)	0.495 (0.464-0.526)	0.570 (0.539-0.601)
xlmr-large	0.718 (0.689-0.746)	0.624 (0.593-0.654)	0.648 (0.617-0.678)	0.508 (0.477-0.539)	0.523 (0.492-0.554)	0.672 (0.642-0.701)	0.500 (0.469-0.531)	0.622 (0.591-0.652)	0.919 (0.900-0.935)	0.925 (0.907-0.941)	0.929 (0.911-0.944)	0.831 (0.806-0.854)	0.846 (0.822-0.868)	0.845 (0.821-0.867)	0.414 (0.383-0.445)	0.480 (0.449-0.511)	0.511 (0.480-0.542)	0.492 (0.461-0.523)	0.442 (0.411-0.473)	0.585 (0.554-0.616)
xlmr-xl	0.850 (0.826-0.872)	0.656 (0.626-0.685)	0.563 (0.532-0.594)	0.659 (0.629-0.688)	0.669 (0.639-0.698)	0.757 (0.729-0.783)	0.525 (0.494-0.556)	0.777 (0.750-0.802)	0.910 (0.891-0.927)	0.934 (0.917-0.949)	0.946 (0.930-0.959)	0.896 (0.875-0.914)	0.904 (0.884-0.922)	0.869 (0.846-0.889)	0.503 (0.472-0.534)	0.568 (0.537-0.599)	0.552 (0.521-0.583)	0.558 (0.527-0.589)	0.478 (0.447-0.509)	0.561 (0.530-0.592)
xlmr-xxl	0.824 (0.799-0.847)	0.743 (0.715-0.770)	0.655 (0.625-0.684)	0.647 (0.616-0.677)	0.647 (0.616-0.677)	0.720 (0.691-0.748)	0.691 (0.661-0.720)	0.687 (0.657-0.716)	0.921 (0.903-0.937)	0.955 (0.940-0.967)	0.949 (0.933-0.962)	0.929 (0.911-0.944)	0.902 (0.882-0.920)	0.931 (0.913-0.946)	0.482 (0.451-0.513)	0.527 (0.496-0.558)	0.495 (0.464-0.526)	0.501 (0.470-0.532)	0.495 (0.464-0.526)	0.505 (0.474-0.536)

Test suites