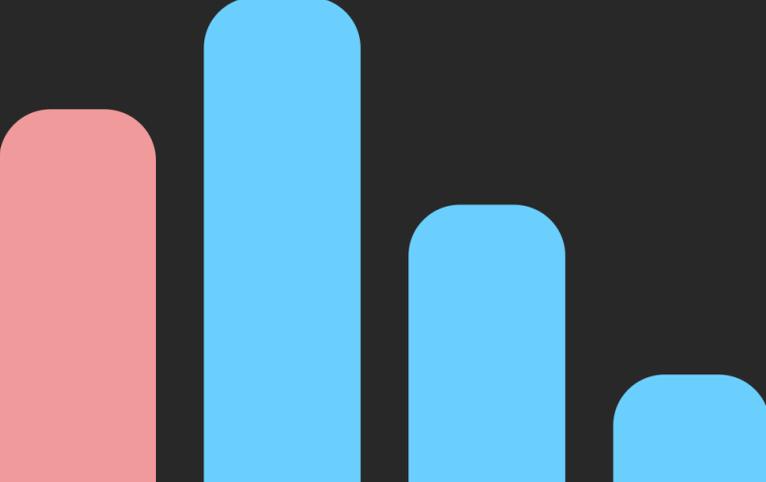


Maturaarbeit

# Abstimmungs- und Wahlvorhersagen

Übertragung amerikanischer Modelle zur Vorhersage und Analyse von Wahlen auf Schweizer Volksabstimmungen



Kantonsschule Frauenfeld

Autor: Andri Hubalek

Fach: Informatik

Betreuung: Matthias Bolli

## Inhaltsverzeichnis

1. Einleitung	S. 2
2. Existierende Methoden der Wahlvorhersage	S. 3
2.1 Ursprünge der umfragebasierten Wahlvorhersagen	S. 3
2.2 Weiterentwicklungen in der Umfragetechnik	S. 4
2.3 Wahlprojektionen basierend auf grundlegenden Indikatoren	S. 7
2.4 Computergestützte Vorhersagen	S. 10
2.5 Fazit aus den historischen Beispielen	S. 11
3. Theoretische Form der App	S. 12
3.1 Grundlegendes & Idee	S. 12
3.2 Erster Teilbereich: Datenbetrachtung	S. 12
3.3 Zweiter Teilbereich: Parolen-Einfluss	S. 13
3.4 Dritter Teilbereich: Das Vorhersagemodell	S. 13
4. Praktische Umsetzung der App	S. 16
4.1 Grundlegende Informationen zum Prozess	S. 16
4.2 Programmierung und Datensammlung für die Datenbetrachtung	S. 17
4.3 Programmierung der Parolen-Analyse-Tools	S. 19
4.4 Erstellung des Modells	S. 20
4.5 Design der App	S. 25
5. Diskussion	S. 26
5.1 Resultate des Modells	S. 26
5.2 Allgemeines Fazit	S. 28
6. Schlusswort	S. 30
6.1 Zusammenfassung	S. 30
6.2 Persönliche Erkenntnisse	S. 31
7. Benutzeranleitung	S. 32
8. Quellen	S. 34
8.1 Primärliteratur und Datenquellen	S. 34
8.2 Sekundärliteratur	S. 34
8.3 Internetartikel	S. 34
8.4 Internetressourcen	S. 35
8.5 Abbildungsverzeichnis	S. 35

## 1. Einleitung

Während des dritten und vierten Jahres an der Kantonsschule Frauenfeld schreiben die Schüler eine Maturaarbeit. Diese sollte eigenständig durchgeführt werden, zur Betreuung steht eine Lehrperson beiseite. Ziel der Arbeit ist die Formulierung einer Fragestellung, sowie die Erarbeitung und Dokumentation ebendieser.

Diese Arbeit setzt sich mit der Vorhersage von Abstimmungen und Wahlen auseinander. Ich bin sowohl an Geschichte und Politik als auch an Mathematik und Informatik sehr interessiert, und dieses Thema erlaubt es mir, beide Teilbereiche zu bearbeiten. Spätestens seit den Präsidentschaftswahlen 2020 in den Vereinigten Staaten fragte ich mich, wieso in der Schweiz nicht auch so viele verschiedene Modelle und Methoden benutzt werden, um mögliche Ausgänge von Abstimmungen zu bestimmen. Hierfür fand ich eine geschichtliche Erarbeitung interessant. Ausserdem konnte ich, da ich ein eigenes Modell als App zur Verfügung stellen würde, Einblicke in die App-Entwicklung gewinnen.

Die zentrale Fragestellung ist, ob und wie die Methoden aus amerikanischen Wahlvorhersagen benutzt werden können, um die Resultate von Schweizer Volksabstimmungen vorherzusagen. Ich habe mehrere Hypothesen aufgestellt. Ich nehme sehr stark an, dass sich die in den Vereinigten Staaten verwendeten Methoden auf viel mehr verschiedene Modelle und Ansätze beziehen als die in der Schweiz verwendeten Umfragen. Dies liegt vor allem am grösseren Interesse am Ausgang von amerikanischen Wahlen. Dieses grössere Interesse ist einerseits politisch; ein neuer Präsident kann in den Vereinigten Staaten einen massgeblichen politischen Kurswechsel auslösen. Andererseits ist auch das mediale Interesse an Wahlen in den Vereinigten Staaten grösser. In der Schweiz existiert vor Wahlen meist weniger Interesse am Ergebnis und deutlich mehr an der Meinungsbildung. Die zweite Hypothese ist, dass sich die amerikanischen Methoden auch auf die Schweiz übertragen lassen. Hier erwarte ich aber auch gewisse Veränderungen. Einerseits, da es sich um verschiedene politische Systeme handelt, andererseits, da in der Schweiz über Vorlagen abgestimmt wird, und in den Vereinigten Staaten über Personen. Die dritte Hypothese ist, dass eine Vorhersagemethode, welche für die Schweiz entwickelt wurde, in der Lage sein sollte, Resultate zu einem zufriedenstellenden Grad vorherzusagen, der den Abweichungen von Umfragen nahe kommen sollte. Diese Vorhersagemethode will ich dann als Applikation einem Benutzer zugänglich machen. Ausserdem sollten auch einige der für die Analyse der amerikanischen Vorhersagemethoden benutzten Daten in der Applikation einsehbar sein. Weil ein grosser Teil der aufgewendeten Arbeit im Fach Informatik angesiedelt ist, ist dies auch eine Arbeit im Fach Informatik. Allerdings setzte ich mich nicht mit einem rein informatischen Thema auseinander. Dementsprechend ist diese Arbeit auch strukturiert. In einem ersten Teil analysiere ich die amerikanischen Wahlvorhersagemethoden. Dieser Teil ist vor allem historisch und war sehr rechercheintensiv. Darauf folgen die Teile, welche sich mit der Umsetzung der Programmierung auseinandersetzen.

Ich danke meinem betreuenden Lehrer, Matthias Bolli, für seine Unterstützung und hilfreichen Inputs. Ausserdem danke ich meiner Mathelehrerin, Hannah Keller, und meinem Geschichtslehrer, Peter Giger, für ihre Beratung und Unterstützung. Ausdrücklich will ich auch Prof. Dr. Sarah Dégallier Rochat danken, welche als Betreuerin für *Schweizer Jugend forscht* massgebliche Inputs gegeben hat, ohne welche die jetzige Form der Arbeit nicht möglich gewesen wäre. Zuletzt danke ich allen, denen ich mit der Ausrede, ich müsse an der Arbeit schreiben, ausgewichen bin, für ihre Geduld und ihr Verständnis.

## 2. Existierende Methoden der Wahlvorhersage

### 2.1 Ursprünge der umfragebasierten Wahlvorhersagen

Die ersten Meinungsumfragen wurden vermutlich in den Vereinigten Staaten durchgeführt. Die erste bekannte Umfrage fand 1824 in Wilmington, Delaware statt.<sup>1</sup> 335 Stimmen gingen an Andrew Jackson, nur 169 an John Quincy Adams.<sup>2</sup> Dies reflektierte die Resultate bei der echten Abstimmung – Jackson gewann eine Pluralität der Stimmen. Derartige Meinungsumfragen wurden immer wieder in verschiedensten amerikanischen Orten veranstaltet, um ein Gefühl für den Ausgang der Wahl zu erhalten. Normalerweise wurden Passanten befragt, weswegen derartige Umfragen oft nur lokal durchgeführt wurden. Dies änderte sich ab 1916 – der in New York ansässige *Literary Digest* führte eine landesweite Befragung durch. In der Ausgabe vom 16. September 1916 befand sich ein Artikel mit einem Schein-Wahlzettel.<sup>3</sup> Der Leser wählte darauf dann den Kandidaten aus, welcher in seiner Gemeinde am populärsten war – sowohl für die anstehende Wahl, als auch für die von 1912. Dies ermöglichte eine rudimentäre Art der Gewichtung. Eine Überrepräsentierung einer gewissen Wählergruppe könnte erkannt werden, wenn die erhaltenen Resultate für die jeweils letzten Wahlen nicht mit den echten Wahlergebnissen übereinstimmten.

Der *Digest* führte seine Umfragen in folgenden Jahren auf ähnliche Art weiter, wechselte allerdings dazu, nach den Präferenzen der Antwortenden, und nicht den Sentiments in ihren Heimatorten, zu fragen. Es stieg auch die Menge der erhaltenen Antworten. Von 30'000 bei der ersten Durchführung 1916 auf über zwei Millionen im letzten Durchführungsjahr 1936. Die Genauigkeit der Vorhersagen war erstaunlich – sie wichen nur wenige Prozentpunkte von den echten Resultaten ab.<sup>4</sup> Allerdings war die Umfrage auch nicht ohne Kritik. 1924 erschien in der *New York Times* zum Beispiel ein Artikel, in welchem dargestellt wurde, dass der *Digest* vor allem wirtschaftlich besser betuchte Bürger befragte.<sup>5</sup> Dies zeigte sich anhand der jeweiligen Vorjahreszahlen, welche stärker in die Richtung der Präferenzen wohlhabender Klassen tendierten, wie dies von einer repräsentativen Bevölkerung zu erwarten wäre. Allerdings konnte auch die Wahl von 1924 korrekt vorhergesagt werden.

Im letzten Durchführungsjahr änderte sich dies. 1936 sagte der *Digest* einen überwältigenden Sieg vorher. Allerdings für den falschen Kandidaten.<sup>6</sup> Präsident Franklin Roosevelt gewann seine zweite Amtszeit – der *Digest* hatte einen klaren Sieg für seinen Gegner Alf Landon vorhergesagt. Dies befeuerte die gleiche Kritik,

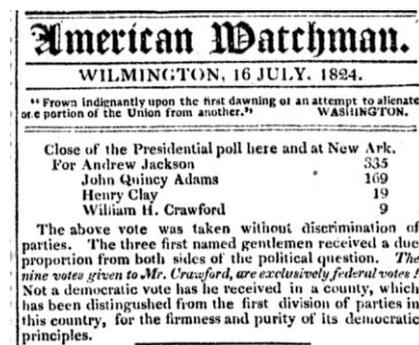


Abb. 1: Ausschnitt aus dem *American Watchman*, welcher die wohl erste Meinungsumfrage durchführte. Erschienen am 16. Juli 1824.

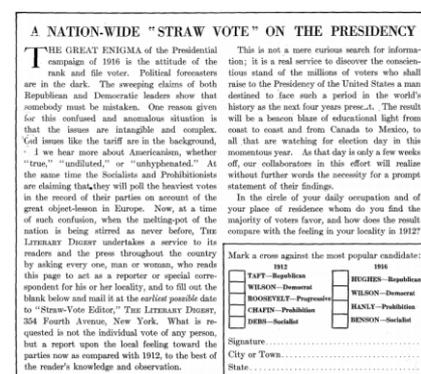


Abb. 2: Artikel im *Literary Digest* des 16. September 1916, in welchem ein Umfragezettel enthalten ist.

<sup>1</sup> Tankard, James W.: *Public Opinion Polling by Newspapers in the Presidential Election Campaign of 1824* in *Journalism Quarterly*, S. 361

<sup>2</sup> *American Watchman* 16. Juli 1824, S. 3

<sup>3</sup> *Literary Digest* 16. September 1916, S. 9

<sup>4</sup> Campbell, W. Joseph: *Lost in a Gallup*, S. 42

<sup>5</sup> *The New York Times* 4. Oktober 1924: *The Literary Digest Poll*, S. 12

<sup>6</sup> *Literary Digest* 31. Oktober 1936, S. 1ff

wie sie schon im Artikel von 1924 laut wurde. Es würden nur Wähler erreicht werden, welche wohlhabender waren, und diese Wähler würden den republikanischen Kandidaten bevorzugen. Allerdings war dies nicht der Ursprung der Fehlvorhersage. Reichere Bürger bevorzugten – wenn auch weniger klar – Roosevelt. Ausserdem zeigten die Zahlen des *Digest* auch für ärmere Bürger eine Präferenz für Landon. Die wahrscheinlichste Antwort liegt in der Polarisierung. Es war die Aufgabe der Befragten, die Postkarten, die der *Digest* verschickte, wieder zu retournieren. Von den 10 Millionen Karten kamen etwa 2,35 Millionen wieder zurück. Da Roosevelt mit vergleichsweise starker Hand regierte, waren seine Gegner entsprechend engagiert gegen ihn. Seine Befürworter waren deutlich weniger motiviert – immerhin hatte Roosevelt die letzten Wahlen mit grossen Mehrheiten gewonnen, man konnte beruhigt sein. Sie antworteten seltener. Für den *Literary Digest* kamen aber Erklärungsversuche zu spät – das Magazin stellte 1938 den Verkauf ein und wurde vom *Time Magazine* übernommen.<sup>7, 8</sup>

## 2.2 Weiterentwicklungen in der Umfragetechnik

1936 war zwar für den *Literary Digest* ein katastrophales Jahr, allerdings war es auch die Geburtsstunde fortgeschrittenerer Umfragemethoden. Die Umfragen des *Digest* hatten keine Gewichtung und keine Kontrolle darüber, wer dem Magazin eine Antwort hatte zukommen lassen, was zur katastrophalen Fehlvorhersage von 1936 geführt hatte. Gleichzeitig etablierten sich aber auch neue Umfrageinstitute, welche wissenschaftliche

Methoden nutzten, um repräsentative Umfragen schaffen zu können. Unter den neuen Organisationen war die von George Gallup gegründete Gallup Organization die wichtigste.<sup>9</sup>

Gallup publizierte schon vor dem *Digest* seine Daten, welche sich auf eine deutlich kleinere Menge befragter Wähler stützten. Hierfür erhielt er einige Kritik, der Redakteur des *Literary Digest* bat darum, dass «Doktor Gallup seine politische Wahrsagerei auf das Büro [seines Instituts] beschränken würde und den *Literary Digest* und seine Zahlen freundlichst und vollständig allein lassen würde.»<sup>10</sup> Auch der republikanische Vize-Parteichef war Gallup gegenüber negativ eingestellt – immerhin sagten Gallups Zahlen die Niederlage seiner Partei vorher.<sup>11</sup>

Entsprechend gross war der Triumph für die wissenschaftlicheren Umfragen. Allerdings ging unter, dass auch diese Vorhersagemethode Roosevelts Mehrheit unterschätzt hatte. Trotzdem gelang es aber, ein deutlich repräsentativeres Bild zu schaffen. Die hierfür angewandte Methodik änderte sich immer wieder, meist

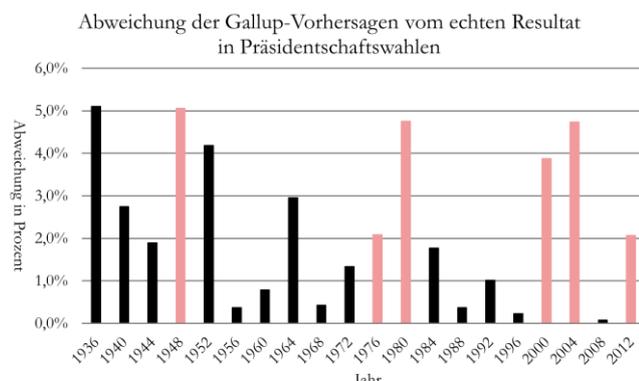


Abb. 3: Abweichungen der Gallup-Vorhersagen von den echten Resultaten. Abweichung ist der Betrag der Differenz des vorhergesagten Prozentsatzes des Gewinners und seines wahren Prozentsatzes. In Rot (■) sind die Vorhersagen, welche den Gewinner nicht richtig identifizierten.

<sup>7</sup> Campbell, W. Joseph: *Lost in a Gallup*, S. 56 - 61

<sup>8</sup> Time Magazine 23. Mai 1938: *Digest Digated*, S. 46 & 47

<sup>9</sup> Campbell, W. Joseph: *Lost in a Gallup*, S. 44

<sup>10</sup> The New York Times 19. Juli 1936: *Dr. Gallup chided by Digest Editor*, S. 21

<sup>11</sup> The New York Times 1. November 1936: *Republican scores Poll*, S. 38

nachdem nach Fehlvorhersagen – siehe hierzu Abb. 3 – Fehlerquellen in den Daten gefunden werden konnten. So funktionierte die Methodik von 1936 auch für Roosevelts erneute Wiederwahlen 1940 und 1944.

1948 sagten dann aber alle drei nationalen Umfrageinstitute, die Gallup Organization sowie die Institute unter Archibald Crossley und Elmo Roper, einen Sieg für den republikanischen Gegner des Präsidenten Harry Trumans vorher. Da diese drei Institute die Vielzahl der Zeitungen belieferten und somit die allgemeine Erwartungshaltung prägten, kam der Sieg Trumans für Zeitungen und Bürger unerwartet.<sup>12</sup> Freude gab es beim bereits erwähnten, nun aber ehemaligen, Redakteur des *Literary Digest*: «Auch wenn es vielleicht falsch ist, so zu reagieren, muss ich doch sagen, dass es mich tröstet, zu wissen, dass andere den gleichen Fehler gemacht haben wie wir.»<sup>13</sup> Die Umfrageinstitute suchten nach ihren Fehlern, ein Komitee wurde einberufen, um den Fehlvorhersagen nachzugehen. Dieses veröffentlichte 1949 seine Resultate, wobei sich herausstellte, dass sich die Ungenauigkeit aus vielen kleineren Fehlern zusammensetzte.<sup>14</sup> Es wurde auf Quoten gesetzt, um eine repräsentative Darstellung der Bevölkerung zu erhalten. Dabei konnte die ausführende Person Antwortende frei wählen, musste schlussendlich aber gewisse Quoten erfüllen.<sup>15</sup> Diese richteten sich nach Geschlecht, Alter, Bildung und Herkunft. Einerseits konnten die implizite Voreingenommenheit des Befragenden eine Auswirkung auf die Auswahl der Antwortenden haben, andererseits richteten sich die Quoten nicht nach der Bevölkerung, sondern nach der Teilmenge der Personen, die in der letzten Abstimmung gewählt hatten. Nach dem Erfüllen dieser Quoten wurden die Personen, die angegeben hatten, nicht wählen zu wollen, aus der Umfrage gestrichen. So wurden Gruppen, welche weniger wählten, unterrepräsentiert. Hiermit wurde deren Nichtwählerrate doppelt eingerechnet. Erhebliche Fehlerquellen waren Unentschiedene sowie Nichtwähler. Nichtwähler konnten nicht sicher aussortiert werden. Zwar wurde nach der Wahlabsicht gefragt, allerdings hielten sich einige nicht an ihre Aussagen. Von den unentschiedenen Wählern stellte sich schliesslich heraus, dass sie sich zu gut  $\frac{3}{4}$  für Truman entschieden hatten. Ausserdem hatten viele Wähler ihre Entscheidung in den letzten zwei Wochen vor der Wahl gefällt. Die Umfrageinstitute hatten zu diesem Zeitpunkt ihre Tätigkeiten bereits eingestellt, immerhin waren die Daten gleichbleibend und klar.

Erste Verbesserungen aus diesen Erkenntnissen führten aber 1952 zu einer weiteren Niederlage.<sup>16</sup> Man meinte, aus den Erfahrungen von 1948 gelernt zu haben. Alle drei wichtigen Institute wollten sich nicht festlegen und sagten, dass es vermutlich knapp werden würde. Ihre Zahlen zeigten dies nur bedingt – Gallups letzte Umfrage vor der Wahl zeigte beispielsweise 47% für Dwight Eisenhower und 40% für seinen demokratischen Gegner Adlai Stevenson.<sup>17</sup> Die unentschiedenen Wähler gewichtete Gallup dann aber im Verhältnis drei zu eins für die Demokraten, da diese Verteilung sich vier Jahre zuvor so gezeigt hatte. Aufgrund dieser Berechnung ergab sich ein Resultat, welches beide Kandidaten gleichauf zeigte. Die aus ähnlichen Methoden entstandenen Fehlvorhersagen aller drei führte zum schlussendlichen Untergang der Crossley-Firma sowie der Nicht-Publikation der Roper-Resultate.

---

<sup>12</sup> The New York Times 4. November 1948: *The Common Man dazed by Election*, S. 8

<sup>13</sup> The Evening Standard 4. November 1948: *Last Editor of Literary Digest Laughs at Error of Other Polls*, S. 6

<sup>14</sup> Wilks, Samuel S. et al: *Report on the Analysis of Pre-Election Polls and Forecasts* in Public Opinion Quarterly, S. 599ff

<sup>15</sup> Time Magazine 3. Mai 1948: *The Black and White Beans*, S. 21 - 23

<sup>16</sup> Campbell, W. Joseph: *Lost in a Gallup*, S. 85 - 107

<sup>17</sup> The Gallup Poll: Public Opinion 1935 – 1971 Vol. 2, S. 1100

Daraufhin wurden die Erkenntnisse aus diesen Niederlagen genutzt, um einige Änderungen an den Umfragen vorzunehmen, welche in den Folgejahren zu erheblichen Verbesserungen der Genauigkeit führten.<sup>18</sup> Unentschiedene Wähler wurden nicht einfach gestrichen, oder wie 1952 nach alten Zahlen verteilt, sondern in die Endberechnung miteinbezogen. Sie wurden gefragt, in welche Richtung sie tendieren, was ein ungefähres Bild der Präferenzen in dieser Gruppe gab und die Zahl der wirklich Unentschiedenen verkleinerte. Ausserdem wurden verbesserte Methoden benutzt, um herauszufinden, ob eine Person wählen wird oder nicht. Mehrere Fragen wurden gestellt. Einerseits waren dies grundlegende Fragen, zum Beispiel, ob die befragte Person wählen wolle und wie sicher diese Intention sei. Andererseits wurde auch gefragt, wo sich das nächstgelegene Wahllokal befinde. Durch Erfahrungswerte konnte dann jedem Antwortgebenden eine Wahlwahrscheinlichkeit zugeschrieben werden. Für die Auswahl der Befragten wurde das alte Quotensystem abgeschafft und durch ein zufälliges und damit repräsentativeres System ausgetauscht. Gallup nutzte ein System der zufälligen Auswahl von Ortschaften, in welcher dann jedes n-te Haus befragt wurde. Hierbei fragte man immer nach dem jüngsten wahlberechtigten Mann, und für den Fall, dass keiner verfügbar war, nach der jüngsten wahlberechtigten Frau. Damit konnten diese sonst unterrepräsentierten Gruppen in einem angemessenen Mass erreicht werden. Junge Menschen und vor allem Männer waren seltener zu Hause anzutreffen, da sie arbeiteten. Mit dem Aufkommen von Telefonen konnten diese genutzt werden, um mögliche Wähler zu befragen. Durch zufälliges Wählen von Nummern konnte eine grosse und repräsentative Teilmenge der Bevölkerung erreicht werden. Diese verbesserte Methodik führten zu einer Art goldenem Zeitalter für Umfragen. Es wurden in den Jahren 1956 bis 1972 einige sehr genaue Vorhersagen getroffen.

1980 folgte wiederum eine sehr grosse Abweichung. Der sehr klare Sieg Ronald Reagans konnte nicht erkannt werden. Wiederum fielen viele Wähler ihre Entscheidung vermutlich erst spät, weswegen ihre Meinungsänderung nicht mehr aufgenommen werden konnten. Hauptgründe für einen Umentschied waren eine schlechte Wirtschaft, sowie die weiterhin anhaltende Geiselnahme amerikanischer Diplomaten in Teheran. Reagan wurde als besser für Wirtschaftswachstum und als stärkerer und entschiedenerer Anführer gesehen.<sup>19</sup> Die spätere Vermutung, dass die frühe Publikation von Nachwahlbefragungen aus den östlichen Staaten Wähler im Westen demotivierte und diese deswegen nicht wählten, konnte nicht belegt werden. Klar war, dass es Menschen im Westen Amerikas störte, dass bereits Sieger erklärt wurden, obwohl in ihren Staaten noch abgestimmt werden konnte. Den grossen Fernsehnetzwerken war es möglich, bereits klare Vorhersagen zu treffen, als die Urnen im Osten zuzogen, was aufgrund der verschiedenen Zeitzonen mehrere Stunden vor Urnenschluss im Westen passierte. Trotz vieler Anekdoten über Menschen, die, nachdem sie von den Resultaten aus dem Osten gehört hatten, die Warteschlange zum Wahlbüro verliessen, konnte bei Nachforschungen nie eine Wirkung auf Wahlergebnisse gemessen werden.<sup>20</sup>

Im 21. Jahrhundert kam es vor allem bei Gallup zu grösseren Fehlern. 2000 wurden sich täglich ändernde Zahlen herausgegeben, welche sehr starke Fluktuationen in der Wählermeinung aufzeigten. Hier schien wiederum das Bemessen der Wahlabsicht das Hauptproblem gewesen zu sein. Es wurde mehrere Wochen vor der Wahl bereits nach Wahlabsichten gefragt, weswegen hier vermutlich nicht die wirkliche Wahlabsicht

---

<sup>18</sup> Perry, Paul.: *Election Survey Procedures of the Gallup Poll* in *Public Opinion Quarterly*, S. 531 - 542

<sup>19</sup> The New York Times 16. November 1980: *Poll Shows Iran and Economy Hurt Carter Among Late-Shifting Voters*, S.1 & 32

<sup>20</sup> Adams, William C.: *Election Night News and Voter Turnout*, S. 111 & 128

am Abstimmungstag, sondern viel mehr der Enthusiasmus für die eine oder andere Partei gemessen wurde. Wenn eine Gruppe motivierter war, gab sie mit grösserer Wahrscheinlichkeit an, wählen zu wollen.<sup>21</sup> Als 2004 und 2012 wiederum falsche Vorhersagen, nicht zuletzt aufgrund veralteter Methodik zur Bestimmung der Wahlwahrscheinlichkeit sowie Sparmassnahmen, folgten, entschied sich die *Gallup Organization* dazu, keine Wahlvorhersagen mehr zu produzieren und sich auf die Marktforschung zu konzentrieren.<sup>22</sup>

Es gibt auch heutzutage viele kleinere und grössere Marktforschungsunternehmen wie auch Zeitungen, welche Umfragen sowohl auf nationaler, als auch auf regionaler Ebene produzieren. Ausserdem haben sich auch alternative Formen von Umfragen etablieren können, die vor allem auf Schwarmintelligenz setzen. Befragte Personen werden nicht nach ihren eigenen Präferenzen gefragt, sondern nach ihrer Einschätzung der Wahlergebnisse. Statt also eine Teilmenge der Wählerschaft nach ihren Absichten zu befragen und dies dann auszuwerten, überlässt man die Auswertung gewissermassen dem Befragten und sammelt dann diese Informationen. Zur Sammlung solcher Sentiments können auch Märkte genutzt werden, in welchen die Mitspieler von marktwirtschaftlichen Anreizen motiviert werden. Vorhersagen mit Schwarmintelligenz sind relativ genau, wobei Märkte der Sammlung der Vorhersagen leicht unterlegen sind.<sup>23</sup> Märkte haben den Vorteil, dass sie häufig bis zum Abstimmungsdatum durchgeführt werden können und wenig Auswertung benötigen. Allerdings haben beide dieser Methoden die Schwäche, dass sie nicht in einem Vakuum existieren. Es ist wahrscheinlich, dass befragte Personen sich bei ihren Einschätzungen auf Daten aus Umfragen stützen und ohne diese wohl nur weniger genaue Angaben machen könnten.<sup>24</sup>

### 2.3 Wahlprojektionen basierend auf grundlegenden Indikatoren

Mit den Gründen für eine Wahlentscheidung, sowohl der Entscheidung, überhaupt zu wählen, setzen sich Umfrageinstitute meist nur auseinander, wenn es darum geht, mögliche Fehlerquellen in Umfragen zu suchen und diese zu korrigieren. Die Frage, was Menschen überhaupt zu ihren Entscheidungen führt, ist aber eigentlich deutlich fundamentaler als das Bestimmen eines Resultates in einem Jahr. Nur schon die Faktoren, welche eine Person dazu bewegen, überhaupt abzustimmen, sind vielfältig, von Dingen wie der grundlegenden politischen Überzeugung, über die demographischen Daten einer Person, bis hin zum Wetter am Wahltag.<sup>25</sup> Modelle, welche auf Indikatoren basieren, setzten sich allerdings nicht mit individuellen Bürgern auseinander, vielmehr analysieren sie Faktoren, welche in der Vergangenheit eine Korrelation mit der Wahlentscheidung gezeigt haben.

Das bekannteste derartige Modell ist das vom Geschichtspräsidentenprofessor Allan J. Lichtman entwickelte *13 Keys*-Modell, mit welchem seit 1984 alle amerikanischen Präsidentschaftswahlen vorhergesagt werden konnten.<sup>26</sup> Hierfür wurden alle Wahlen seit 1860 analysiert, um die wichtigsten Faktoren zu finden, welche für das Resultat entscheidend waren. Grundannahme des Systems ist, dass ein Präsident, beziehungsweise der Kandidat seiner Partei, wiedergewählt wird. Sind allerdings genügend viele der 13 Indikatoren falsch, also gegen

---

<sup>21</sup> Erikson, Robert S. et al: *Likely (and unlikely) voters and the assessment of campaign dynamics* in Public Opinion Quarterly, S. 595

<sup>22</sup> Enten, Harry: *Gallup Gave Up. Here's Why That Sucks*

<sup>23</sup> Atanasov, Pavel et al: *Distilling the Wisdom of Crowds: Prediction Markets vs. Prediction Polls*, S. 14 & 15

<sup>24</sup> Erikson, Robert S. et al: *Are Political Markets Really Superior to Polls as Election Predictors?* in Public Opinion Quarterly, S. 210

<sup>25</sup> Adams, William C.: *Election Night News and Voter Turnout*, S. 14 - 20

<sup>26</sup> Lichtman, Allan J.: *Predicting the next President*, S.197

den Amtsinhaber, sagt das Modell vorher, dass die Bevölkerung den Präsidenten ersetzen wird, da zu viel Unzufriedenheit mit der regierenden Partei vorhanden ist. Da die USA gleichzeitig immer von nur zwei dominanten und nicht grundlegend ideologisch geprägten Parteien regiert wurde, gab es meistens nur eine Alternative zu der regierenden Partei.<sup>27</sup> Unter gewissen Annahmen kann auch mathematisch gezeigt werden, dass das System zu einem Zweiparteiensystem tendiert.<sup>28</sup> Da sich beide Parteien um die politische Mitte streiten, kommt es immer wieder zu Machtwechseln, welche relativ wenig mit den zugrundeliegenden politischen Zielen der Parteien zu tun haben.

Im *13-Keys*-Modell sind alle 13 Indikatoren gleich stark gewichtet. Interessanterweise kann so jede Wahl genau vorhergesagt werden, allerdings korrelieren die Werte nur bedingt mit den erhaltenen Wählerstimmen ( $|r| = 0,68$ ). Dies kann durch alternative Gewichtung der Indikatoren verbessert werden. Erlaubt man zum Beispiel fünf verschiedene Gewichtungsstufen (0,2, 0,4, 0,6, 0,8 und 1,0), so kann der Korrelationskoeffizient gesteigert werden ( $|r| = 0,81$ ), ohne die Richtigkeit der Vorhersage zu ändern. Dies gibt zwar Aufschluss über die relative Stärke der Indikatoren, garantiert aber nicht, dass die Vorhersagen wirklich richtig bleiben. Eine optimierte Korrelation setzt nicht die Richtigkeit aller Vorhersagen voraus, da eine kleine Verschiebung in der Indikatorenge weighting die eine Vorhersage zwar falsch machen kann, allerdings die Korrelation mehrerer anderer Vorhersagen verbessern kann, was den gesamtheitlichen Korrelationskoeffizienten steigert. Maximal kann die Korrelation auf  $|r| = 0,83$  gesteigert werden, allerdings entstehen dann bereits falsche Resultate. Jeder Indikator ist mindestens einmal relevant – es gibt für jeden Faktor eine Situation, in welcher er entscheidend ist, weswegen kein Indikator einfach weggelassen werden kann. Auch die stärker gewichteten Faktoren sind nicht zwingend wichtiger – sie haben lediglich eine grössere Auswirkungen auf die Anzahl Wählerstimmen, die ein Kandidat erhalten wird. Gerade der dritte-Partei-Key hat in seiner Definition schon die Begründung für seine starke Auswirkung auf die Stimmenprozente – er wird erst aktiv, wenn mindestens fünf Prozent der Wähler bewegt werden. In den Jahren, in welchen dies der Fall war, war die Auswirkung auch häufig deutlich grösser als nur fünf Prozent. Gerade das Jahr 1912 sticht heraus, Theodore Roosevelt erreichte mit seiner unabhängigen Kandidatur 27% der Stimmen und zerstörte damit Howard Tafts Chancen auf eine Wiederwahl.<sup>29</sup>

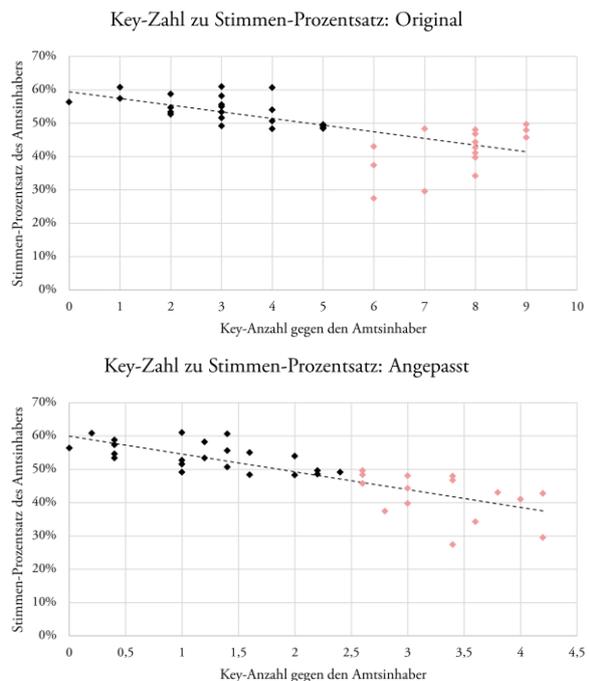


Abb. 4: Originale Key-Anzahl nach Allan J. Lichtmans *Predicting the Next President* oben, eine auf erhöhte Korrelation angepasste Variante unten. Im originalen Modell ist eine Wiederwahl nicht zu erwarten, wenn mehr wie 5 Indikatoren gegen den Amtsinhaber sind. Im angepassten Modell darf das gesamte Gewicht der Indikatoren nicht grösser wie 2,4 sein. In Rot (■) sind alle Jahre, in welchen der Amtsinhaber die Wiederwahl verlor. Dargestellt sind alle Präsidentschaftswahlen von 1860 bis 2008, wobei die Wahlen von 2012, 2016 und 2020 als Kontrolle für das angepasste Modell verwendet wurden.

<sup>27</sup> Chambers, William Nisbet: *Party Development and the American Mainstream in The American Party System*, S. 3 & 28

<sup>28</sup> Palfrey, Thomas R.: *A Mathematical Proof of Duverger's Law in Models of Strategic Choice in Politics*, S. 83

<sup>29</sup> The San Francisco Call 6. November 1912: *Maelstrom of Votes leaves Bull Moose and G.O.P. wrecked*, S. 1

**TABELLE 1**

SCHLÜSSEL / INDIKATOREN NACH ALLAN J. LIGHTMANS <i>PREDICTING THE NEXT PRESIDENT</i>				
<i>Nummer</i>	<i>Name</i>	<i>Beschreibung</i>	<i>Gewichtung Original</i>	<i>Gewichtung Angepasst</i>
1.	Partei-Mandat	Nicht erfüllt, wenn die Amtsinhaber-Partei nach den Midterm-Wahlen weniger Sitze im Kongress hat wie zuvor.	1,0	0,2
2.	Kein Konflikt um Nominierung	Nicht erfüllt, wenn es einen innerparteilichen Konflikt um die Nominierung des Kandidaten in der Amtsinhaberpartei gibt.	1,0	1,0
3.	Amtsinhaber tritt an	Erfüllt, falls der aktuelle Amtsinhaber und nicht nur ein Parteigenosse antritt.	1,0	0,2
4.	Keine dritte Partei	Erfüllt, wenn keine dritte Partei mit mehr als 5% der zu erwartenden Stimmen antritt.	1,0	1,0
5.	Kurzfristige Wirtschaft	Erfüllt, wenn sich die Wirtschaft während der Wahl nicht in einer Rezession befindet.	1,0	0,4
6.	Langfristige Wirtschaft	Erfüllt, wenn Wirtschaftswachstum höher wie in zwei vorhergehenden Amtszeiten ist.	1,0	0,2
7.	Wichtige Gesetzesänderung	Erfüllt, wenn eine signifikante Gesetzesänderung bewirkt wurde.	1,0	0,4
8.	Keine sozialen Unruhen	Nicht erfüllt, wenn es während der Amtszeit anhaltende soziale Unruhen gab.	1,0	0,8
9.	Keine Skandale	Nicht erfüllt, wenn der Amtsinhaber in einen grossen Skandal involviert war.	1,0	0,2
10.	Keine Rückschläge in der Aussenpolitik	Nicht erfüllt, wenn grosse Rückschläge oder Imageschäden in der Aussenpolitik aufgetreten sind.	1,0	0,2
11.	Erfolg in der Aussenpolitik	Erfüllt, wenn der Amtsinhaber einen grossen Erfolg in der Aussenpolitik zu verbuchen hat.	1,0	0,2
12.	Amtsinhaber-Charisma	Erfüllt, wenn der Amtsinhaber besonders charismatisch oder ein Nationalheld ist.	1,0	0,6
13.	Kein Herausforderer-Charisma	Nicht erfüllt, wenn der Herausforderer besonders charismatisch oder ein Nationalheld ist.	1,0	0,2

Neben diesen relativ offensichtlichen Erklärungen zeichnen sich allerdings doch auch Muster ab: Die Schlüssel zwei, vier und acht sind am aussagekräftigsten. Alle sind Anzeichen für eine unruhige Amtszeit und Unzufriedenheit sowohl in der Bevölkerung als auch in der eigenen Partei. Dies wiederum scheint die Theorie des Systems, dass ein Präsident wiedergewählt wird, ausser er macht einige Fehler oder ist zu passiv, zu bestätigen.

Schlussendlich sind die *Keys* aber weniger ein genaues Vorhersagemodell und mehr ein formalisiertes Modell zur Analyse wichtiger politischer Faktoren und Einflüsse. Für die Wahl von 2016 wurde das System angepasst<sup>30</sup> – es sagte nun nicht mehr den Gewinner im Sinne von Stimmenprozenten, sondern den Gewinner des *Electoral College* vorher. Dies, da in den letzten Jahren im stark demokratischen Staat Kalifornien Millionen von Stimmen für den demokratischen Kandidaten abgegeben wurden, welche im *Electoral College* und deswegen für das Resultat der Wahl allerdings keine Veränderung bewirkten. Der Gewinner eines Staates erhält alle Elektorenstimmen dieses Staates, unabhängig davon, wie gross die Gewinnmarge war. Nur diese Änderung erlaubte die korrekte Vorhersage des Wahlsieges Donald Trumps, welcher gut 3 Millionen weniger Stimmen wie seine Konkurrentin Hillary Clinton erhielt. Ein System, welches auf allgemeinen Faktoren basiert, ist von Veränderungen in diesen Faktoren abhängig und muss von Experten gegebenenfalls angepasst werden.

Beispielsweise suggeriert das *Keys*-System, das der Wahlkampf eines Kandidaten keinen Einfluss auf das Resultat hat. Allerdings könnte dies auch nur daran liegen, dass normalerweise beide Kandidaten einen relativ ähnlichen Wahlkampf führen. Sollte sich ein Kandidat entscheiden, keinen aktiven Wahlkampf zu führen, keine Interviews zu geben und sich wenig öffentlich zu zeigen, könnte sich daraus ein Effekt auf die Resultate ergeben, was dann eine Anpassung des Vorhersagesystems bedingen würde. Systeme, die auf

<sup>30</sup> Lichtman, Allan J.: *Predicting the next President*, S.191

historischen Begebenheiten beruhen, können sich schwer tun, sobald unkonventionelle Methoden genutzt werden. Da es sich nur um vereinfachte Versionen komplexerer Analysen handelt, müssen Systeme, die auf externen Faktoren beruhen, auch anpassbar sein, sollten sich neue Begebenheiten ergeben.

## 2.4 Computergestützte Vorhersagen

Mit der verstärkten Verbreitung von Computern wurde es möglich, Daten schneller zu sammeln, auszuwerten und zu publizieren. Dies wiederum führte dazu, dass neue Arten entwickelt wurden, um die bereits existierenden Daten zu verwerten. Einerseits können Berechnungen angestellt werden, welche aus früheren Tendenzen einzelner Staatsbevölkerungen die Präferenzen und die Verteilung der möglichen Resultate in diesen Staaten berechnen. Andererseits werden auch Umfragedaten auf Staatsebene, welche wiederum aufgrund der Entwicklung des Internets einfacher gesammelt werden können, für die Berechnung der möglichen Resultate in einem Staat verwendet. Schlussendlich ergibt sich eine Verteilung möglicher Resultate, welche die jeweilige Wahrscheinlichkeit wiedergibt.<sup>31</sup> Um dann ein gesamtes Resultat zu erhalten, werden zufällig Resultate – entsprechend ihrer Wahrscheinlichkeit – ausgewählt. Diese können dann in einer Monte-Carlo-Simulation zusammengefasst werden. Bei einer solchen Simulation werden viele zufällig ausgewählte mögliche Resultate erstellt und dann gesammelt, um einen Mittelwert zu errechnen. Dies ist sinnvoll, da nicht etwa ein System analysiert wird, in welchem der Kandidat mit den meisten Stimmen gewinnt, sondern eines, in welchem die Mehrheiten in den einzelnen Staaten relevant sind, vergleichbar mit dem Schweizer Ständemehr. Dabei ist wichtig, dass die zufällig ausgewählten Ereignisse auch in sich stimmig sind. Wird für einen Staat zum Beispiel ein unwahrscheinlich gutes Resultat für einen Kandidaten zufällig ausgewählt, so sollte dies die Auswahl eines ähnlichen Resultates in einem ähnlichen Staat wahrscheinlicher machen.

Beispielsweise räumte das Modell von *fvethirtyeight.com*, welches erstmals 2008 benutzt wurde und seitdem im Feld der computergestützten Vorhersagemodelle eine führende Position eingenommen hat, Donald Trump 2016 eine Gewinnchance von 30% ein. Zwar war diese Zahl vermutlich etwas zu tief, immerhin gewann Trump die Wahl, allerdings war es doch eine sehr hohe Chance, verglichen mit Umfragen, konventioneller Weisheiten und anderen Modellen.<sup>32</sup> Die meisten anderen Vorhersagen des Volksmehrs waren auch relativ genau.<sup>33</sup> Sie vernachlässigten aber die Möglichkeit, dass ein Kandidat mit einer Minderheit der Stimmen das *Electoral College* gewinnt. Da Trump gleich in mehreren Staaten im *Rust Belt* knappe und unwahrscheinliche Siege erzielte – motiviert von ähnlichen Sentiments – konnte er ebendies knapp bewerkstelligen. Da diese Staaten im *fvethirtyeight*-Modell aufgrund ihrer Ähnlichkeit aneinander gekoppelt waren, konnte ein Resultat, in welchem sie alle ähnlich agierten, mit grösserer Wahrscheinlichkeit vorhergesagt werden. Dieser Ähnlichkeitsfaktor, die Kovarianz, kann unter anderem aus demographischen Daten sowie historischen Abstimmungsmustern errechnet werden.

Speziell an diesen Modellen ist, dass in ihnen ersichtlich ist, dass es sich nur um Vorhersagen handelt, welche eintreffen könnten, und welchen gewisse Wahrscheinlichkeiten zugeschrieben werden können. Es wird keine eindeutige Vorhersage gemacht. Dies ist aufgrund der meist unklaren Daten sowie der durch die Natur

---

<sup>31</sup> Rigdon, Steven E. et al: *A Bayesian Prediction Model for the U.S. Presidential Election* in American Politics Research, S. 700ff

<sup>32</sup> Katz, Josh: *Who Will Be President?*

<sup>33</sup> Campbell, W. Joseph: *Lost in a Gallup*, S. 184

der Statistik gegebenen Unsicherheiten auch oftmals der beste Schluss aus den vorhandenen Informationen. Nur schon an der Fehlerrate der normalen Umfragen – siehe Abb. 3 – zeigt sich, dass es besser ist, Unsicherheiten zu kommunizieren und mit ihnen zu rechnen. Einige moderne Modelle rechnen auch mit grundlegenden Faktoren, um Unsicherheit zu messen und entsprechend anzupassen. So können beispielsweise Wirtschaftsindikatoren oder die Menge an *New York Times*-Titeln, welche sich über die gesamte Breite der Frontseite erstreckten, was normalerweise wichtigen Nachrichten vorbehalten ist, Indikatoren für eine Unsicherheit in der Bevölkerung und die daraus entstehenden neuen und unberechenbaren politischen Entscheidungsfaktoren sein.<sup>34</sup> Allgemein können fast alle Datenpunkte, welche eine sinnvolle Korrelation mit Wahlergebnissen zeigen, benutzt werden, allerdings ist es wichtig, ihren Einfluss richtig zu messen und zu gewichten.

## 2.5 Fazit aus den historischen Beispielen

Aus der Geschichte der amerikanischen Wahlvorhersagen zeigt sich vor allem eines: Es handelt sich um eine sehr ungenaue Wissenschaft, welche sich mit der Zeit aber um einiges verbessert hat. Scheinbar befindet sich die Vorhersagetechnik immer auf einer Art Gratwanderung zwischen verblüffend genauen Vorhersagen und kompletten Fehltreffern. Dieses Muster zeigte sich sowohl beim *Literary Digest*, welcher sich vor seiner Fehlvorhersage 1936 noch selbst für seine Genauigkeit lobte, wie auch bei den modernen Methoden von *FiveThirtyEight.com*, welche 2008 und 2012 praktisch perfekte Resultate produzierten, und 2016 Hillary Clinton eine 71%-Chance zuschrieben, die Wahl zu gewinnen.<sup>35</sup> Auch die Methoden, welche nicht auf Umfragen basieren, sind nicht ganz sicher vor Fehlern. Wie bereits erwähnt sind sie nur eine Vereinfachung der von Experten beobachteten Indikatoren und müssen deswegen gegebenenfalls neuen politischen Realitäten angepasst werden. Insgesamt wird für die Entwicklung eines eigenen Modells also vor allem wichtig sein, nicht zu viel zu erwarten, vor allem, weil sehr viele mögliche Parameter unbekannt sind. Dies ist gerade auch eine Stärke der neusten Computervorhersagen, welche bereits mit Unsicherheiten rechnen. Ziel eines selbst programmierten Modells sollte also auch sein, diese Unsicherheiten darzustellen. Egal, was die Vorhersagemethode war, Übermut war oftmals das grösste Problem vieler vorhersagender Politikwissenschaftler.

---

<sup>34</sup> Silver, Nate: *How FiveThirtyEight's 2020 Presidential Forecast Works – And What's Different Because of Covid-19*

<sup>35</sup> Thomas, Andrew: *538's Uncertainty Estimates Are As Good As They Get*

### 3. Theoretische Form der App

#### 3.1 Grundlegendes & Idee

Am Anfang der Arbeit stand selbstverständlich die Frage, was überhaupt das Thema werden würde. Der ungefähre Rahmen war schon früh klar, gerade aber die Funktionen der App entwickelten sich während der Ausarbeitung der Arbeit laufend. Schon anfangs war mir klar, dass die verwendete Sprache C# sein würde. Die andere Programmiersprache, die ich an der Kantonsschule gelernt hatte, war Python. Python eignet sich nur bedingt zum Schreiben von Anwendungen mit einem graphischen Benutzerinterface. Mit C# nutzte ich erst Windows Forms und wechselte anschliessend auf Xamarin. Xamarin erlaubt die Bereitstellung der gleichen App auf iOS, Android und Windows. Da das iPad an der Kantonsschule als Arbeitsgerät genutzt wird, schien es mir sinnvoll, eine App für das iPad zu entwickeln.

Diese App ist in drei Bereiche unterteilt. Der erste Teil ist schlicht ein Betrachter für viele der Basisdaten, welche im ersten Teil verwendet wurden. Der zweite Teil erlaubt eine rudimentäre Analyse der Einflüsse von Parteiparolen. Dieser bildet eine Grundlage für den dritten Teil, welcher ein Vorhersagemodell beinhaltet, welches sich auf grundlegende Faktoren stützt. Hierfür ist relevant, inwieweit sich amerikanische Analysemethoden auf das Schweizer System übertragen lassen. Zwar hat die Schweiz ein anderes politisches System, und gerade beim Vergleich von Präsidentschaftswahlen und Volksabstimmungen gibt es betreffend Meinungsbildung und Wirkung von Persönlichkeiten einige Unterschiede. Allerdings sind auch einige Parallelen sichtbar. Die Schweiz hat mehrere Umfrageinstitute, welche Meinungsbefragungen durchführen. Diese sind meist genau, aber vor Fehlern auch nicht sicher. Eine grosse Unsicherheitsquelle sind, wie auch in Amerika, Unentschiedene, allerdings sind sie normalerweise in kleineren Zahlen vertreten als in den Vereinigten Staaten. Auch hierzulande sind Umfragen nicht sicher vor Anschuldigungen, Wahlergebnisse zu beeinflussen.<sup>36</sup> Diese Anschuldigung wird vor allem dadurch vorgebeugt, dass zwei Wochen vor dem Abstimmungssonntag keine neuen Umfragen mehr publiziert werden. Dies führt aber auch dazu, dass die neusten Umfragen zum Zeitpunkt einer Abstimmung mindestens zwei Wochen alt sind. Die Meinungsbildung kann in dieser Zeit noch fortschreiten und zu Abweichungen zwischen den beiden Momentaufnahmen Umfrage und Abstimmung sorgen.

Gerade deswegen wollte ich versuchen, ein Modell zu programmieren, welches ohne Umfragedaten auskommt. Im Teilkapitel zu den *13 Keys* wurde bereits thematisiert, dass Indikatoren genutzt werden können, um eine relativ genaue Vorstellung vom Resultat einer Wahl zu erhalten. Es ist zu erwarten, dass relativ gute Resultate gefunden werden können, wenn die Indikatoren genutzt werden, um möglichst ähnliche Abstimmungen zu finden und mit deren Resultaten zu rechnen.

#### 3.2 Erster Teilbereich: Datenbetrachtung

Diese Funktionalität lässt sich in zwei Bereiche unterteilen: Einerseits ist dies die Datenbeschaffung, andererseits die Programmierung der Funktionalität. Für die Schweizer Volksabstimmungen stehen mehrere einheitliche Datensätze zur Verfügung, für die amerikanischen Umfragen und Resultate war mehr Recherche nötig. Nachdem diese Daten gesammelt – manchmal auch aus Primärquellen abgetippt – waren, mussten

---

<sup>36</sup> Senti, Martin: *Wetterleuchten der Demoskopie in Wahl-Probleme der Demokratie*, S. 76

sie vereinheitlicht werden. Im Hintergrund ist jeder Datenpunkt als eine Datei abgelegt. Ein Problem stellen die verschiedenen Formen der Daten dar, die zur Verfügung standen. Während für Abstimmungsresultate Daten auf Kantons- beziehungsweise Staatsebene existieren, bestehen Umfragen häufig nur aus einem nationalen Prozentsatz. Hierfür sind verschiedene Anzeigeoptionen nötig.

Für Daten auf regionaler Ebene kann neben dem Gesamtergebnis auch eine Karte angezeigt werden. Für Umfragen oder die *13 Keys* ist eine solche Darstellung nicht möglich. Ihre Visualisierung besteht aus deutlich simpleren Balkendiagrammen beziehungsweise Listen der Keys. Für die genaueren Daten können aber zusätzliche Funktionen genutzt werden. Beispielsweise können Ergebnisse verschoben werden, so dass ein Resultat angezeigt werden kann, welches entstanden wäre, wenn die eine oder andere Seite einige Prozentpunkte mehr oder weniger erreicht hätte. Bei schlichten Umfragen ist diese Auswertung trivial und deswegen nicht Bestandteil der App. Wenn allerdings die Mehrheiten von und in Staaten relevant sind, ist diese Betrachtung wertvoll. Auch ist es sinnvoll, diese komplexeren Auswertungen untereinander vergleichen zu können und diesen Vergleich zu visualisieren.

### 3.3 Zweiter Teilbereich: Parolen-Einfluss

Diese Funktion dient vor allem als Vorbereitung auf den dritten Teil. Es handelt sich um eine kleine Funktion, welche für jede Partei, jede Institution oder jeden Verband mit herausgegebenen Parolen angibt, wie häufig die Parolen mit dem Abstimmungsresultat übereinstimmen. Ausserdem können dann einzelne Organisationen kombiniert werden. Zum Beispiel könnte analysiert werden, wie häufig eine Parole, welche die FDP herausgibt, auch angenommen wird. Es wäre aber auch möglich, zu sehen, wie häufig eine Parole, die vom Bund angenommen, aber beispielsweise von der SVP abgelehnt wird, durchkommt. Diese Kombinationsmöglichkeit ist essenziell, da die Resultate deutlich mehr Aussagekraft haben, umso mehr Parteien sie einschliessen. Es ist zu erwarten, dass die Mitteparteien die grössten Erfolgsprozentsätze haben. Dies liegt aber nicht daran, dass diese Parteien alleine ausschlaggebend sind. Vielmehr ist es sehr häufig, dass sie mit mindestens einer Randpartei übereinstimmen, und die kombinierten Wählerschaften dann reichen, um eine Mehrheit zu erzielen. Der Erfolgsprozentsatz ist also weniger ein Indikator dafür, was die Partei selbst bewirkt, sondern viel mehr ein Indikator dafür, wie mehrheitsfähig Vorlagen mit der Zustimmung einer gewissen Partei sind. Auch viele Vereine und Organisationen haben bereits Parolen herausgegeben. Während die Anzahl dieser Parolen oftmals deutlich zu klein ist, um eine wirkliche Aussagekraft zu haben, kann es trotzdem interessant sein herauszufinden, wie sich ein Verein in der Vergangenheit engagiert hat.

### 3.4 Dritter Teilbereich: Das Vorhersagemodell

Ein Modell kann auf vielen Indikatoren basieren. Einerseits könnte man sich auf Umfragen beziehen. Diese ziehen einige Probleme mit sich. Einerseits sind sie alles andere als unfehlbar; das Institut gfs.bern, welches Abstimmungsumfragen für die SRG ausführt, zeigt in gut 80% der Fälle das richtige Resultat, verfehlt aber häufig die genauen prozentualen Verteilungen.<sup>37</sup> Ein darauf basierendes Modell hätte die gleichen Fehler und wäre kaum in der Lage, weitere Informationen beizutragen. Andererseits ist gerade bei Volksinitiativen immer auch das Ständemehr relevant – Umfragedaten aus den Kantonen fehlen aber. Basierend auf einer

---

<sup>37</sup> Kovic, Marko: *Longchamps SRG-Umfragen im 10-Jahres-Check*

landesweiten Umfrage könnten so schwer Projektionen für einzelne Kantone erstellt werden. Es sollte allerdings möglich sein, aufgrund einiger externer Faktoren ähnliche Volksabstimmungen in der Schweizer Geschichte zu finden und aufgrund derer Resultate Projektionen zu erstellen.

Nach Claude Longchamp, dem Gründer von gfs.bern, lassen sich Volksabstimmungen in gewisse Muster einordnen.<sup>38</sup> Das Resultat ist abhängig von mehreren Faktoren, einerseits von der Grundeinstellung der Bevölkerung, andererseits aber auch der Meinungsbildung im Wahlkampf. Zu unterscheiden ist auch zwischen Initiativen und Referenden. Bei Ersteren handelt es sich meist um grundsätzliche Entscheidungen. Dies liegt in der Natur der Initiative als Verfassungsänderung. Deswegen können die meisten Stimmberechtigten schon aufgrund ihrer grundlegenden politischen Meinung schnell ein Urteil fällen. Unter Volksinitiativen zeigen sich meist zwei Typen. Einerseits gibt es Anliegen von Minderheiten, welche meist klar abgelehnt werden und meist mit dem Ziel initiiert werden, eine Diskussion anzuregen. Andererseits gibt es auch Volksinitiativen, welche möglicherweise mehrheitsfähig sind. Hier zeigt sich anfangs oft eine starke Zustimmung, welche dann im Verlauf des Abstimmungskampfes abnimmt. Auch diese Anliegen werden meist abgelehnt. Deutlich bessere Chancen haben normalerweise Gesetze, gegen welche ein Referendum ergriffen wird. Am häufigsten angenommen werden die Gesetze, zu welchen schon vor dem Abstimmungskampf klare positive Meinungen existierten. Allerdings kommt es bei Referenden auch häufig vor, dass anfangs die meisten Stimmbürger unentschieden sind, sodass der Abstimmungskampf eine entscheidende Rolle spielt.

Ein Modell sollte in der Lage sein, derartige Muster zu erkennen und ähnliche Vorlagen zu finden. Hierfür sind vier verschiedene Indikatoren sinnvoll. Erstens sollte die Rechtsform der Vorlage angegeben werden. Die Unterscheidung zwischen Initiative und Referendum ist stark ausschlaggebend. Volksinitiativen werden zu gut 90% abgelehnt. Im Gegensatz dazu wird der Bund bei Referenden meist bestätigt. Das Resultat einer Volksinitiative anhand der Resultate ähnlicher Referenden zu bestimmen, würde höchstwahrscheinlich eine deutlich zu optimistische Projektion ergeben. Der zweite wichtige Indikator ist der Politikbereich. Frühere Resultate zu Vorlagen, die beispielsweise die direkten Steuern betreffen, sind für eine neue Vorlage, welche auch die direkten Steuern betrifft, deutlich aussagekräftiger als eine Vorlage, die zum Beispiel die Aussenpolitik betrifft. Zur Einordnung wird die dreistufige Einteilung nach dem Swissovotes-Datensatz verwendet.<sup>39</sup> Dieser unterscheidet erst grössere Bereiche, wie beispielsweise Wirtschaft oder Energie, und unterteilt diese dann weiter. So ist beispielsweise Arbeit und Beschäftigung ein Unterthema des Bereiches Wirtschaft. Diese Unterthemen haben meist weitere Teilbereiche. So lässt sich die Ähnlichkeit eines Bereiches zu einem anderen abstufen, abhängig davon, ob sie den exakt gleichen Teilbereich, oder nur das gleiche grössere Thema betreffen. Das Thema selbst sagt aber noch nichts darüber aus, ob eine Vorlage beispielsweise für oder gegen den Ausbau der Atomkraft ist. Hierfür sind Parolen hilfreich. Mit ihnen lässt sich zeigen, wie die Parteientiments verteilt sind. Sinnvoll ist auch, die jeweiligen Parolen in der Ähnlichkeitsberechnung entsprechend den Parteistärken zu gewichten. Eine abweichende Parole der Mitte ist deutlich wichtiger als eine der EVP. Die vierte relevante Angabe ist auch die simpelste. Seit Beginn der Volksabstimmungen hat sich sowohl die Parteilandschaft als auch die Zusammensetzung des Stimmvolkes verändert. Deswegen ist

---

<sup>38</sup> Longchamp, Claude: *Prädisposition bei Schweizer Volksabstimmungen in Direkte Demokratie: Forschung und Perspektiven*, S. 300 & 301

<sup>39</sup> Ein Link zum Datensatz ist unter 8.4 in den Quellen zu finden

es sinnvoll, das gegebene Jahr als zusätzlichen Indikator zu verwenden. Der Benutzer gibt dann alle Indikatoren der Vorlage an, die vorherzusagen sind. Der Vergleich der historischen Indikatoren jeder Vorlage mit den vom Benutzer angegebenen Indikatoren kann dann dazu genutzt werden, um jeder historischen Abstimmung einen Ähnlichkeitswert zuzuschreiben. Pro Kanton können dann mit diesen Ähnlichkeiten sowie den jeweiligen Abstimmungsergebnissen Mittelwerte und Standardabweichungen berechnet werden. Der Mittelwert ist hierbei das wahrscheinlichste Resultat in einem gegebenen Kanton anhand der vom Benutzer gewählten Indikatoren.

Basierend darauf können Simulationen durchgeführt werden. Hierfür werden alle kantonalen Verteilungen zu einer multivariaten Normalverteilung zusammengerechnet. Zu dieser Verteilung gehören ein Mittelwertvektor, welcher mit den Mittelwerten aller kantonalen Verteilungen befüllt wird, sowie eine Kovarianzmatrix, welche alle Kovarianzen zwischen den Kantonen beinhaltet. Diese Kovarianzen werden anhand der bereits vergangenen Abstimmungen berechnet, geben also an, wie die Abstimmungen in den Kantonen zusammenhängen. Aus diesen Werten kann die multivariate Verteilung erzeugt werden, woraus dann wiederum einzelne Punkte auf dieser Verteilung herausgezogen werden können. Jeder dieser Punkte ist ein mögliches Resultat, ein mögliches Szenario, eine Simulation.

Nach Abschluss der Simulation können relevante Informationen angezeigt werden, wie zum Beispiel der bereits genannte Mittelwert, oder der Anteil der Simulationen, in welchen die Vorlage angenommen werden würde. Bei Volksabstimmungen sind auch Angaben wie beispielsweise der Mittelwert der Anzahl zustimmender Stände relevant. Ausserdem sollte auch immer eine Art Vertrauensindikator angegeben werden. Dieser sollte einerseits bereits in die Standardabweichungen eingeflossen sein sowie aber auch separat angegeben werden. Der Indikator sollte aufgrund der Anzahl sehr ähnlicher gefundener historischer Resultate berechnet werden. Ist diese Zahl höher, stützt sich die Vorhersage auf mehr Daten und ist damit vertrauenswürdiger.

## 4. Praktische Umsetzung der App

### 4.1 Grundlegende Informationen zum Prozess

Im Informatikunterricht lernten wir den Umgang mit C# und Benutzeroberflächen vor allem mit der Benutzung von Windows Forms. Mit Windows Forms lassen sich simple Windows-Applikationen erstellen. Es sind einige grundlegende Steuerelemente vorhanden. Da ich mich mit diesen bereits auskannte, wollte ich erst auch Windows Forms nutzen, um mein Programm zu präsentieren. Allerdings haben diese Applikationen immer einige Probleme. Konkret war das grösste aber die Verfügbarkeit. Zwar ist es eigentlich relativ einfach, eine exe-Datei aus dem Internet herunterzuladen, aber gerade in der Schule kann dies umständlich sein, da installierte Programme nicht von einer Arbeitsstation zu einer anderen übertragen werden können. Das Programm könnte auch auf dem individuellen Schülerordner abgelegt werden, allerdings ist der dort vorhandene Platz stark begrenzt. Mein Betreuer schlug vor, das Projekt auf eine Webseite umzulagern. Ich versuchte dies, allerdings bemerkte ich, dass mir diese Art der Programmierung noch ein wenig zu unbekannt war. Ausserdem hätte ich mich um Hosting und den damit verbundenen Aufwand kümmern müssen. Wäre das Programm auf einem Server ausgeführt worden, hätte dies ausserdem nochmal einen grösseren Aufwand verursacht. Die clientseitige Ausführung hätte wohl Leistungsprobleme mit sich gezogen. Deswegen entschied ich mich für die Nutzung von Xamarin. Der Download aus dem App-Store ist simpel, Verfügbarkeit war also kein Problem mehr. Allerdings musste ich mich nun um die entsprechende Apple-Entwickler-Lizenz kümmern. Ausserdem können iOS-Apps nur auf MacOS kompiliert werden. Deswegen musste mein Windows-Desktop jeweils per SSH mit einem MacBook verbunden werden. Glücklicherweise übernimmt die Visual-Studio-IDE hier die Konfiguration. Auch das Debug-iPad musste bei Apple speziell registriert werden. Als dies dann getan war, lief der Entwicklungsprozess relativ unkompliziert.

Jede der Funktionen besteht aus einer eigenen Seite. Diese wiederum besteht prinzipiell aus zwei Dokumenten. Eines gibt die Struktur der Benutzeroberfläche wieder, das andere definiert die Funktionen, welche ausgeführt werden, wenn der Benutzer mit den Steuerelementen interagiert. Zusätzlich hierzu können weitere Code-Dokumente benutzt werden, um gewisse Prozesse und Funktionen auszulagern. Normalerweise werden hier bestimmte Muster verwendet. So sollte beispielsweise der Code, der für die Benutzeroberfläche zuständig ist, separat von dem Code gespeichert werden, der für die grundlegende Funktionsweise des Programms zuständig ist. Ich habe mich im Verlauf der Arbeit aber nur bedingt an diese Idealform gehalten.

Zusätzlich zu diesen funktionsgebenden Dokumenten arbeitet die App auch mit Dokumenten, welche zur Datenspeicherung genutzt werden. Diese konnte ich teilweise aus grösseren Dokumenten generieren, manchmal musste ich sie aber auch von Hand schreiben. Ich arbeitete mit Tab-Separated-Values, kurz tsv. Dies erlaubt die Speicherung von Daten in einer Tabellenstruktur. Spalten werden durch einen Tab getrennt, Zeilen durch einen Zeilenumbruch. Dies hat den Vorteil, dass der Text auch ohne speziellen Viewer relativ gut als Tabelle erkennbar ist. Das war wiederum wichtig, da ich Microsoft Excel, wenn überhaupt, nur vorsichtig benutzen konnte, um Daten zu übertragen. Excel hat die Tendenz, Kommazahlen automatisch und irreversibel zu Kalenderdaten umzuwandeln. Wenn meine App dann eine Kommazahl erwartet und beispielsweise «25. Oktober» zurückkommt, folgt verständlicherweise ein Error.

Alle diese Dateien, also Code und Daten beziehungsweise Ressourcen, lud ich periodisch auf die Plattform Github.<sup>40</sup> Diese erlaubt es, den Code auf mehreren Geräten herunterzuladen und zu synchronisieren. Ausserdem ist er so öffentlich und kann eingesehen werden. Als die App dann fertig war, konnte ich sie kompilieren und in den Apple App-Store hochladen.

#### 4.2 Programmierung und Datensammlung für die Datenbetrachtung

Den ersten Teil der App, den Datenbetrachter, erstellte ich noch, bevor ich das Projekt auf Xamarin übertragen hatte. Ausserdem wusste ich dort noch nicht genau, wie Code auf mehrere Dokumente verteilt wird, weswegen der ganze Code in einem einzigen Dokument zu finden ist. Dies ist unübersichtlich, der Code generell ist aufgrund mehrerer Überarbeitungen nur schwer lesbar.

Grundsätzlich sind aber vor allem die Dokumente im Hintergrund wichtig. Der Benutzer hat zuerst die Wahl zwischen den beiden Ländern, den Vereinigten Staaten und der Schweiz, die ich in der Arbeit betrachtet habe. Vor allem in der Sektion der Vereinigten Staaten sind hier interessante Daten. Es existieren Datensätze zu den Präsidentschaftswahlen, konkret die echten Resultate, die Vorhersagen des *Literary Digest*, die der *Gallup Organization* und die Stellungen der *13 Keys*. Wie bereits erwähnt waren verschiedene Grade an Datenqualität verfügbar.

Die besten Daten waren die offiziellen Resultate. Diese sind ausführlich dokumentiert und geben für jeden Staat und Kandidaten einen genauen Prozentsatz an. Auch sind diese Zahlen im Internet breit verfügbar, weswegen ich mir hier erlaubte, sie von Wikipedia zu übertragen. Die Beschaffung der Resultate der *13 Keys* gestaltete sich auch relativ einfach. Professor Lichtman fasste sie in seinem Buch zusammen. In meiner Ausgabe fehlten allerdings die Angaben für das Jahr 2016. Auch diese Vorhersagen sind aber relativ leicht auf diversen Internetseiten auffindbar, konkret auf der Webseite der American University, an welcher Lichtman unterrichtet. Der *Literary Digest* publizierte seine endgültigen Resultate jeweils kurz vor den Wahlen. Auf der Webseite der amerikanischen *Library of Congress* können für die Jahre 1932 und 1936 andere Magazine gefunden werden, in welchen die Resultate des Digest erneut publiziert wurden. Auf archive.org konnte ich die Ausgaben von 1916 und 1920 finden. Die Jahre 1924 und 1928 konnte ich auf legitimen Wegen nicht finden. Schliesslich konnte ich die eingescannten Dokumente auf einer Webseite aufspüren, welche sich als «alternatives Medium» bezeichnet und Holocaustleugnung betreibt. Sie ist deswegen hier nicht namentlich erwähnt. Ich habe die entsprechenden Scans auf archive.org hochgeladen, wo sie nun verfügbar sein sollten. Die *Digest*-Daten sind immer aufgeteilt nach einzelnen Staaten verfügbar, nur wurden 1916 und 1920 nicht alle Staaten erhoben. Zuletzt blieben noch die Daten von Gallup. Für die Jahre zwischen 1935 und 1977 publizierte die *Gallup Organization* einen Gesamtindex, in welchem die Resultate aller durchgeführter

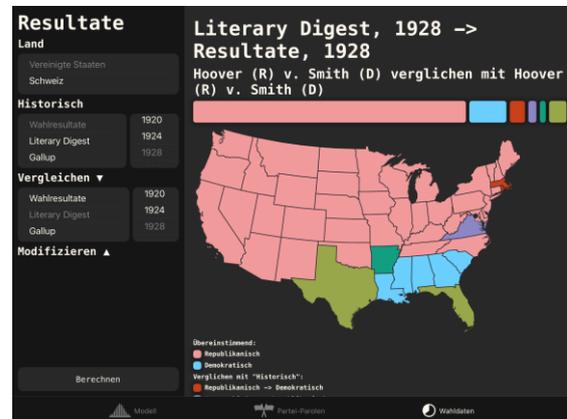


Abb. 5: Screenshot aus dem Datenbetrachter. Hier wurde ein Vergleich zwischen den Daten des *Literary Digest* und den echten Wahldaten für das Jahr 1928 ausgeführt. In Rot und Blau sind die richtig vorhergesagten Staaten, alle andern Farben zeigen eine Abweichung zwischen den beiden Datenpunkten an. Das dunkle Orange zeigt beispielsweise an, dass der *Digest* den betreffenden Staat als republikanisch eingeteilt hat, er in Realität aber für die Demokraten stimmte.

<sup>40</sup> Der Link ist unter 8.4 in den Quellen auffindbar.

Umfragen aufgelistet waren.<sup>41</sup> Auch diese sind digital auf archive.org verfügbar. Für alle folgenden Jahre konnte ich nur über *Roper iPoll* auf die Daten zugreifen. Hierbei handelt es sich um einen bezahlten Datensatz, zu welchem ich mir über die ETH-Bibliothek Zugriff verschaffen konnte. Die dort vorhandenen Daten deckten die restlichen Jahre von 1978 bis 2012 ab. Gallup publizierte nur in frühen Jahren genaue Daten über die einzelnen Staaten, später nur noch nationale Prozentsätze. Manchmal wurden hier noch zusätzlich spezielle Gruppen aufgezeigt. Beispielsweise wurden für die Wahl John Kennedys, des ersten katholischen Präsidenten, die Wahlabsichten nach Katholiken und Protestanten unterteilt. Alle diese Daten übertrug ich dann in einheitliches Format, welches von der App gelesen werden konnte. Die Daten zu Schweizer Volksabstimmungen sind in einem einheitlichen Datensatz von Swissvotes verfügbar und konnten deswegen ohne weitere Komplikationen mit einem Skript in das App-Format übertragen werden.

Der Benutzer kann über ein Menu eine der Dateien auswählen, erst eine Kategorie, also beispielsweise die Resultate des *Digest*, und kann dann ein spezifisches Jahr auswählen. Zu diesem Zeitpunkt liest die App erstmals das Dokument. In der ersten Zeile ist vermerkt, um welchen Datentypus es sich handelt. Dabei sind prinzipiell drei verschiedene Typen möglich. Der Erste stellt eine Karte dar, der Zweite ein Balkendiagramm nach Umfragewerten und der Dritte schlicht eine Liste von Informationen mit Ja/Nein-Angaben. Die erste Variante wird benutzt, wenn Daten auf Staatsebene vorhanden sind. Die Zweite ist für nationale Umfragen und die Dritte für die *13 Keys*. In jedem Dokument folgt der Typusdeklaration in der zweiten Zeile einige Metadaten, beispielsweise die anzuzeigenden Titel und Untertitel. Ausserdem sind die Dateinamen für Farbgradienten und gegebenenfalls Karten angegeben. Darauf folgt eine Zeile mit allen möglichen Parteienamen und darauf eine weitere Zeile mit den beiden wichtigsten Parteien. Der bereits erwähnte Farbgradient ist eine Bilddatei mit einer einzelnen Reihe an Pixeln. Daraus entnimmt das Programm die anzuzeigenden Parteifarben. Diese entsprechen den möglichen Konfigurationen auf der dritten Zeile. Beispielsweise würde diese Zeile mit «Republikanisch» beginnen. Sinnvollerweise wäre der erste Pixel auf dem Farbgradienten rot, die Farbe der Republikaner. Für Umfragen werden knappe Werte mit eigenen Kategorien und Farben dargestellt. Die Kategorie «Tendenz Republikanisch» wird beispielsweise mit einem helleren rot dargestellt. Deswegen sind die beiden Hauptparteinamen auf der vierten Zeile relevant. Hiermit bleibt für das Programm identifizierbar, zu welcher übergeordneten Partei die einzelnen Kategorien gehören.

Auf die ersten vier Metadaten-Zeilen folgen die Datenzeilen, in welchen jeweils ein Datenpunkt angegeben ist. Für Typus eins handelt es sich hierbei um jeweils einen Staat, dessen Gewichtung im *Electoral College* und Gewinner sowie den zugehörigen Prozentsätzen. Es wäre dann also angegeben, dass die Republikanische Partei beispielsweise den Staat Alabama gewonnen hat. Das Programm arbeitet sich durch alle Staaten durch. Hierbei werden einerseits die Stimmen der jeweiligen Kategorien zusammengezählt – also die Anzahl der Stimmen der Republikaner, der Demokraten und möglicher weiterer Parteien. Beim Durcharbeiten durch die Staatenliste wird ausserdem noch die Karte entsprechend eingefärbt. Die Karte ist als Vektor gespeichert. Jeder Staat ist also in der Bilddatei als eigene Form – mit eigener ID – gespeichert. Mit dieser ID kann das Objekt gefunden werden und farblich verändert werden. Die Vektorgrafik kann dann in eine Bitmap, ein Bild, welches auch angezeigt werden kann, umgewandelt werden. Hierfür wird die *SkiaSharp*-

---

<sup>41</sup> The Gallup Poll: Public Opinion 1935 – 1971 Vol. 1 – 3

Package verwendet, welche von Microsoft zur Verfügung gestellt wird. Wenn eine Datei ausgewählt ist, zu welcher diese Kartendaten verfügbar sind, erscheinen zwei weitere Optionen: Modifizieren und Vergleichen. Die Modifizierungsfunktion ist simpel. Der Benutzer gibt einen Stimmen-Prozentsatz an, welcher der einen oder anderen Partei zusätzlich zugeschrieben werden soll. Anhand davon kann bestimmt werden, ob sich die Resultate auf Staatsebene ändern würden. Diese veränderten Ergebnisse können dann, wie bereits ausgeführt, tabelliert werden. Die etwas komplexere Funktion ist die Vergleichsfunktion. Hierbei kann ein Benutzer einen zweiten Datenpunkt auswählen, der verglichen werden soll. Für alle Staaten wird dann betrachtet, ob die Resultate aus den beiden Datensätzen übereinstimmen. So könnte zum Beispiel verglichen werden, inwieweit die Gallup Vorhersagen des Jahres 1936 mit den echten Resultaten übereinstimmen. Die übereinstimmenden Resultate werden bei ihren Farben belassen, für abweichende Resultate, wenn ein Staat beispielsweise für die Demokraten vorhergesagt wurde und von den Republikanern gewonnen wurde, werden weitere zufällige Farben verwendet.

Diese Funktionen, sowohl Modifikation als auch Vergleich, sind für die beiden anderen Darstellungsarten nicht verfügbar. Die beiden anderen Darstellungsarten zeigen lediglich den Inhalt ihrer Datenzeilen. Für die Umfragen bestehen diese Datenzeilen aus der gestellten Frage und dem jeder Partei zugeordneten Prozentsatz. So könnte zum Beispiel zu der Frage der Wahlabsicht die Zahl 0.6 für die Demokraten und 0.4 für die Republikaner vermerkt sein. Die App zeigt dann einen entsprechend gefärbten Balken an, welcher zu 60% das demokratische Blau und zu 40% das republikanische Rot trägt. Die Datenzeilen für die *13 Keys* sind noch simpler, hier ist lediglich die jeweilige Frage und deren Antwort, also wahr oder falsch, aufgeführt. Die App rechnet dann nur kurz zusammen, wie viele Indikatoren positiv und wie viele negativ sind. Dies wird wiederum auf einem Balkendiagramm dargestellt.

Ganz allgemein litt dieser ganze Funktionsbereich vor allem darunter, dass er laufend mit meinen Recherchen entwickelt wurde. Deswegen wurden immer wieder Funktionen hinzugefügt. Dies verursachte ein sehr grosses und unübersichtliches Code-Dokument.

#### **4.3 Programmierung des Parolen-Analyse-Tools**

Die Programmierung dieses zweiten Teils begann erst deutlich später, nach den Sommerferien, nachdem der erste Teil der Arbeit geschrieben und die Recherche abgeschlossen war. Deswegen lief auch die Programmierung deutlich geplanter, der Code ist übersichtlicher. Ausserdem nutzte ich hier auch Klassen. Ich lagerte also Funktionen in andere Dokumente und Funktionsbereiche aus. Dies machte das Programm deutlich übersichtlicher. Neben dem normal generierten Code-Dokument erstellte ich zwei weitere; Vote und Party. Hierbei repräsentiert eine Instanz der Vote-Klasse jeweils eine Volksabstimmung. Beim Start der Applikation wird ein Dokument geladen, in welchem alle Parteien, Organisationen und Institutionen abgelegt sind, die zur Analyse benutzt werden sollen können. Jede dieser Organisationen hat wiederum ein eigenes Dokument, in welchem alle herausgegebenen Abstimmungsparolen aufgeführt sind. Für jede dieser Organisationen wird eine Instanz der Party-Klasse erstellt. Diese enthält eine List von Vote-Klassen. Jede Instanz einer Vote Klasse repräsentiert eine Volksabstimmung und enthält Angaben wie den Namen, den Index, die Resultate und die jeweiligen Parolen. Alle diese Daten stammen aus dem Swissovot-Datensatz,

welcher im Internet frei verfügbar ist. In der App werden alle Organisationen auf einer Seitenleiste dargestellt. Wenn sie ausgewählt werden, erscheint ein Feld für die Partei auf der Hauptbenutzeroberfläche. Hier kann gewählt werden, welche Parteiparolen in die Berechnungen einfließen sollen. So könnte beispielsweise die FDP ausgewählt werden und für sie die Ja-Parole eingestellt werden, sowie die SP mit der Nein-Parole. Das Programm wertet dann alle Abstimmungen aus, in welchen die FDP ja und die SP nein gesagt hat. Diese Wahl läuft über eine Schaltfläche, über welche eine Ja, Nein, oder neutrale Parole ausgewählt werden kann. Sobald eine Auswahl getroffen wird, kann diese genutzt werden, um für jede Partei eine Liste von Vote-Klassen-Instanzen

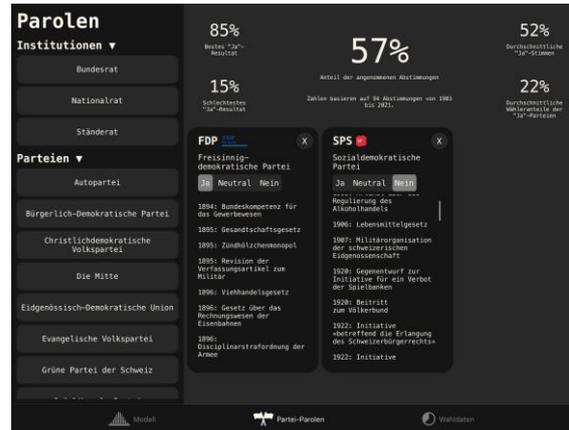


Abb. 6: Screenshot aus der Parolen-Analyse. Hier wird die Abstimmungen analysiert, welche von der FDP befürwortet und von der SP abgelehnt wurden. Die Liste auf den Partei-Steuerenlementen geben alle Abstimmungen an, die für diese Partei analysiert wurden. Im Falle der FDP wäre dies die Liste aller Abstimmungen, welche die FDP befürwortet hat. Die obigen Kennzahlen gelten für die kombinierte Liste aus FDP-Zustimmungen und SP-Ablehnungen. 57% der derartigen Abstimmungen wurden angenommen, mit einem durchschnittlichen Ja-Anteil von 52%.

zu erstellen, welche alle diese bestimmte Parole haben. Diese Liste wäre für das bereits genannte Beispiel also eine Liste aller Vote-Instanzen, zu welchen die Parole der FDP ein Ja war, sowie eine weitere Liste mit allen Vote-Instanzen, zu denen die Parole der SP ein Nein war. Aus den gemeinsamen Elementen wird dann eine gemeinsame Liste erstellt. Diese wird ausgewertet, und Kennzahlen, wie der Prozentsatz dieser Abstimmungen, die angenommen wurden, oder auch der durchschnittliche Ja-Prozentsatz werden berechnet. Die Geschwindigkeit des iPads ist genügend hoch, um diese Berechnungen sofort auszuführen, und nicht wie bei beiden anderen Funktionsteilbereichen einen Knopf zu benötigen. Der Code ist deutlich simpler, als er es noch für den ersten Teil war – allerdings ist die gesamte Funktion auch deutlich weniger kompliziert. Die übersichtlichere Programmierung war auch für den letzten und dritten Teil hilfreich, da dieser erheblich komplexer war.

#### 4.4 Erstellung des Modells

Das Modell basiert darauf, aus Angaben einer vorherzusagenden Abstimmung mögliche Resultate zu erstellen, welche dann ausgewertet werden. Dafür werden alle vorherigen Abstimmungen mit der Vorherzusagenden verglichen. Es wird ein Ähnlichkeitswert bestimmt, welcher dann als Gewichtung benutzt wird, wenn für jeden Kanton ein vermutliches Resultat berechnet wird. Diese vermutlichen Resultate, welche in jedem Kanton als Normalverteilung modelliert werden, können dann zu einer multivariaten Normalverteilung zusammengefasst werden. Aus dieser Verteilung können dann «Stichproben» entnommen werden, wobei jede dieser Proben einem möglichen Resultat entspricht.



Abb. 7: Screenshot aus dem Vorhersagemodell. Die hier gezeigten Resultate entsprechen der Vorhersage zum PMT-Gesetz. Die Graphik zeigt an, wie häufig ein Resultat mit einem bestimmten Resultat berechnet wurde. In den Kenndaten ist der Mittelwert des Ja-Anteils zu finden. Zuletzt folgt eine Liste der Kantone und deren Mittelwerte. Hierfür werden wiederum, wie schon im ersten Funktionsbereich, Vektorbilder der Kantone benutzt, um deren Farben schnell ändern zu können. Die Farben zeigen an, wie sicher das Modell sich ist, dass ein Kanton zustimmt, beziehungsweise ablehnt.

Die Konstruktion des Modells bedingt einige grundlegende Annahmen. Das Modell wird unter Annahme dieser Hypothesen konstruiert. Sie lauten:

1. Frühere Abstimmungen besitzen eine Aussagekraft über kommende Abstimmungen. Wenn zwei exakt identische Abstimmungen stattfinden, werden die Resultate identisch sein. Diese Annahme ist sehr breit gefasst, sollte aber richtig sein, wenn das Stimmvolk nicht zufällig abstimmt. Wenn alle Faktoren, die jemanden zu einer Entscheidung bringen können, identisch sind, sollte die resultierende Entscheidung identisch sein. Ausserdem wird angenommen, dass Vorlagen, welche sich ähnlich sind, ähnliche Resultate haben werden.
2. Es ist bis zu einem gewissen Grad möglich, über Daten zu erfassen, wie ähnlich sich zwei Abstimmungen sind. Hierfür sind das Jahr der Abstimmung, der Politikbereich, die Rechtsform sowie die Parolen der Parteien brauchbare Datenpunkte (Auch wie diese Punkte verglichen werden sind in sich Annahmen, welche in den folgenden Paragraphen noch beschrieben werden). Selbstverständlich sind diese Faktoren alles andere als vollständig, allerdings sollten sie die meisten wichtigen Aspekte abdecken. So sollte das Jahr als Indikator verhindern, dass übliche Abstimmungsmuster, welche sich schon lange nicht mehr so bestätigen, nicht mehr berücksichtigt werden. Rechtsform und Politikbereich sollen zu einer grundsätzlichen Einordnung dienen, sie stellen die Frage, auf wen das Volk bei einem bestimmten Sachverhalt hört. Die Partei-Parolen sollen die politische «Richtung» einer Vorlage erkennbar machen.
3. Um die einzelnen Faktoren untereinander zu gewichten, wird die Konfiguration von Gewichten benutzt, welche für eine frühere Periode die durchschnittlich besten Resultate erzielt. Damit wird angenommen, dass ebendiese Gewicht und Faktoren über diese Zeitperiode gleichbleibend sind, also weiterhin eine Aussagekraft haben. Für das Modell wurde die Periode von 1997 bis 2019 benutzt; die Informationen zu anderen Basis-Datenmengen finden sich in Abschnitt 5.1 *Resultate des Modells*.
4. Die nach Ähnlichkeit gewichteten Resultate können zu Mittelwerten verrechnet werden, welche die Resultate in den jeweiligen Kantonen repräsentieren. Diese Annahme resultiert aus den vorherigen Annahmen.
5. Wie sich die Kantone gegenseitig beeinflussen, ist schlicht von vorherigen Abstimmungen abhängig und unabhängig von Gewichtungen. Die Berechnung der Kovarianzen folgt also der üblichen und nicht einer gewichteten Formel. Diese Annahme basiert darauf, dass die Zusammenhänge zwischen den Verhalten von Kantonen unabhängig von den spezifischen Resultaten sein sollten. Appenzell Inner- und Ausserrhoden stimmen in den meisten Fällen identisch ab, egal, ob es sich um Finanz- oder Sozialpolitik handelt, und unabhängig davon, ob sie die Vorlage ablehnen oder annehmen werden.
6. Die kantonalen Mittelwerte, multipliziert mit der erwarteten Stimmenmenge, und die Varianzen können zu einer multivariaten Normalverteilung zusammengesetzt werden, von der dann zufällige Punkte entnommen werden können, welche jeweils ein mögliches Resultat darstellen. Diese

zufälligen Resultate können in genügend grossen Mengen dafür benutzt werden, um wiederum einen Durchschnitt von Resultaten zu berechnen, welcher als wahrscheinliches Resultat, also die Vorhersage des Modells, interpretiert werden kann.

Am Anfang stehen die Daten, welche im Hintergrund in einem Dokument abgelegt sind. Diese Daten sind mit einem einfachen Skript dem Swissovotes-Datensatz entnommen worden. Zentral ist, dass alle Daten vorhanden sind, die im Modell benutzt werden. In diesem Fall handelt es sich also für jede vergangene Abstimmung um die Indikatoren, welche mit den Benutzereingaben verglichen werden. Diese sind also in diesem Fall das Jahr der Abstimmung, die Rechtsform, der Politikbereich und die Parteiparolen. Ausserdem muss für jeden Kanton sowohl das Resultat in dem Kanton als auch die Anzahl der Abstimmenden vorhanden sein, um kantonale Resultate zu berechnen und zu gewichten.

Im nächsten Schritt wird für jede vergangene Abstimmung ein Ähnlichkeit zur vorherzusagenden Vorlage bestimmt. Dies funktioniert über eine Formel, welche alle Indikatoren beinhaltet. Dabei erhält jeder Indikator ein Gewicht und die finale Summe wird mit einer Aktivierungsfunktion verfeinert. Formeln und Gewichte sind im Diagramm auf der letzten Seite dieses Unterkapitels sichtbar. Für das Jahr wurde eine lineare Funktion benutzt. Sind die Jahre der vorherzusagenden und stattgefundenen Vorlage identisch, ist wird 100% des Gewichts zurückgegeben, sind sie so weit auseinander, wie nur möglich (In diesem Fall 156 Jahre, da schon so lange Abstimmungen durchgeführt werden), ist die Ähnlichkeit des Jahres 0. Für Politikbereich und Rechtsform wird jeweils das Gewicht zurückgegeben, wenn die Indikatoren identisch sind, andernfalls beträgt die Ähnlichkeit 0. Für die Parolen-Ähnlichkeit wird eine Summe berechnet, welche um den Stimmenanteil der Partei grösser wird, wenn die Parolen identisch sind, und um diesen Anteil kleiner wird, wenn die Parolen unterschiedlich sind. Es ergibt sich ein Wert zwischen -1 und 1. Dieser wird dann mit dem der Parolen-Ähnlichkeit zugeschriebenem Gewicht multipliziert. Alle diese vier Indikatoren-Ähnlichkeiten werden addiert und werden zum Input einer Sigmoid-Funktion. Diese Art von Funktion eignet sich gut als Aktivierungsfunktion. Sie besitzt zwei Parameter, die Streckung  $k$  und die x-Achsen-Verschiebung  $x_0$ .

$$\text{Ähnlichkeit} = \frac{1}{1 + e^{-k * (\text{Ähnlichkeit-Jahr} + \text{Ä.-Politikbereich} + \text{Ä.-Rechtsform} + \text{Ä.-Parolen} - x_0)}}$$

Alle diese Gewichte wurden nicht arbiträr gewählt, sie wurden anhand der bereits vergangen Abstimmungen zwischen den Jahren 1997 und 2019 so berechnet, dass für diese Abstimmungen die Abweichungen zwischen Vorhersage und echtem Resultat so klein wie möglich wird. Die hierfür verwendete Excel-Tabelle befindet sich auf Github. Die Optimierungsfunktion ist Teil der Excel Add-Ins Solver und sollte in der Lage sein, ein globales Minimum zu finden, es kann aber nicht ausgeschlossen werden, dass bessere Gewichte möglich sind. Um zu überprüfen, ob diese Minimierungsfunktion nicht einfach gute Resultate für die Abstimmungen von 1997 bis 2019 gefunden hat, welche keine Anwendbarkeit auf andere Vorhersagen haben, wurden die berechneten Gewichte benutzt, um Vorhersagen für alle Vorlagen zu machen, welche ab und mit 2020 stattgefunden haben. Dabei stellt sich heraus, dass die durchschnittliche Abweichung in dieser Kontrollmenge sogar tiefer ist (Die durchschnittliche Abweichung in den Jahren 1997 bis 2019 beträgt 8.62%, die durchschnittliche Abweichung in der Kontrollmenge beträgt 4.88%. Ausserdem fällt auf, dass die Median-Abweichungswerte in beiden Fällen tiefer sind wie die Durchschnittsabweichungen. Für die

Vorhersagen zwischen 1997 und 2019 beträgt die Median-Abweichung 6.88%, in der Kontrolle sind es 2.47%. Dies liegt daran, dass es in der Kontrolle fast keine grossen Ausreisser gibt).

Wenn für alle vergangenen Abstimmungen die Ähnlichkeitswerte berechnet sind, wird für jeden Kanton ein zu erwartendes Resultat ausgerechnet. Dies ist aus zwei Gründen sinnvoll. Einerseits werden kantonale Resultate benötigt, um herauszufinden, ob eine Initiative angenommen werden würde. Das Ständemehr ist relevant für das Resultat der Abstimmung. Andererseits schafft es einen gewissermassen höheren Realismus und verschönert die Darstellung. Dies sind aber sehr oberflächliche Gründe; es wäre also durchaus denkbar, den Schritt mit den Kantonen wegzulassen und einfach eine einzige gewichtete Normalverteilung zu erzeugen, wenn die Resultate auf tieferer Ebene nicht ausschlaggebend sind. Wollte man beispielsweise kantonale Abstimmungen erzeugen, wäre es nicht nötig, Resultate für jede Gemeinde zu berechnen. Wenn dies aber nötig ist, werden in den Kantonen schlicht die gewichteten Normalverteilungen berechnet, wobei die Variable der Prozentsatz Ja-Stimmen einer bestimmten Vorlage in einem Kanton ist und das Gewicht die berechnete Ähnlichkeit dieser Vorlage. Auf gleiche Weise kann auch die erwartete Zahl der Stimmen in einem Kanton berechnet werden, welche als Gewicht wichtig ist. Wenn nur eine Berechnung gemacht würde, könnte nun auch noch die gewichtete Standardabweichung berechnet werden und das Resultat wäre fertig. Da allerdings mehrere Kantone benutzt werden, muss eine multivariate Normalverteilung benutzt werden.

Diese definiert sich einerseits aus einem Vektor bestehend aus den eben berechneten kantonalen Mittelwerten als auch einer Kovarianzmatrix. Um diese Matrix zu befüllen, muss die Kovarianz jedes Kantons zu jedem anderen Kanton berechnet werden. Dabei wird die Gesamtmenge aller bereits stattgefundenen Abstimmungen miteinberechnet. Die Werte ändern sich aber kaum, wenn nur eine Teilmenge, beispielweise die fünfzig letzten Abstimmungen, analysiert werden. Aus Mittelwertvektor und Kovarianzmatrix kann die Verteilung erzeugt werden. Es ist dann möglich, zufällige Punkte auf dieser Verteilung auszuwählen, wobei jeder Punkt einem möglichen Resultat entspricht. Um sicher alle Möglichkeiten abzudecken, werden in diesem Fall eine Million dieser Punkte ausgewertet. Aus jedem dieser Punkte kann ein Gesamtergebn berechnet werden, indem die Kantone anhand ihrer erwarteten Anzahl Stimmen gewichtet werden. Der Durchschnitt dieser Gesamtergebnisse ist dann die Vorhersage des Modells. Diese ergeben wiederum eine Normalverteilung.

Die Resultate werden fertig ausgewertet in der App angezeigt. Es erscheint eine Anzeige mit der Verteilung der Resultate, einige Kenndaten werden darunter angezeigt. Diese Anzeige enthält unter anderem auch den genauen Mittelwert in Prozent, welcher die Vorhersage des Modells darstellt, als auch Informationen dazu, wie viele der Simulationen in einer angenommenen Vorlage resultierten. Ausserdem wird bei Abstimmungen, bei welchen das Ständemehr relevant ist, angegeben, was die durchschnittliche Anzahl Kantone war, welche die Vorlage befürworteten.

Auf der folgenden Seite ist ein Diagramm des Modells mit seinen Parametern und Gewichten (optimiert für die Periode 1997 - 2019) ersichtlich:

Benutzereingabe (Indikatoren der zu machenden Vorhersage):

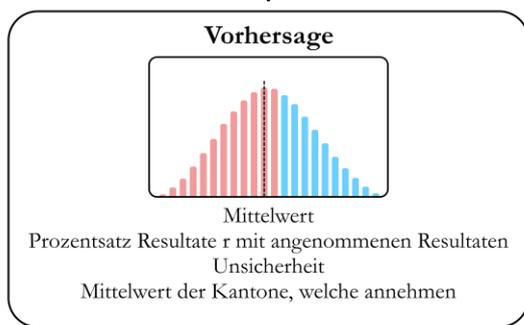
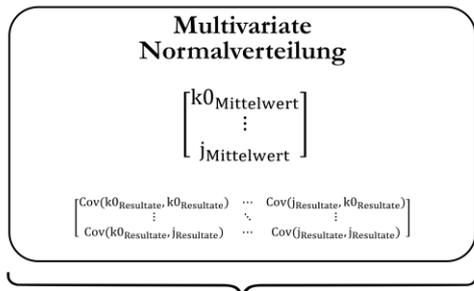
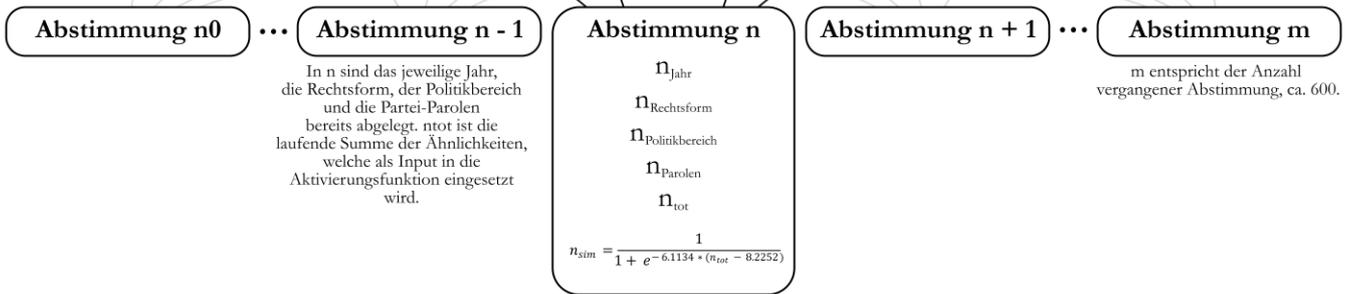


$$n_{tot} += 2.1803 * \left(1 - \frac{|Jahr - n_{Jahr}|}{|2022 - 1866|}\right)$$

$$\text{Wenn } n_{Rechtsform} = \text{Rechtsform, dann } n_{tot} += 0.2702$$

$$\text{Wenn } n_{Politikbereich} = \text{Politikbereich, dann } n_{tot} += 0.5885$$

$$n_{tot} += 2.5932 * \sum_{i=Parolen} (\text{Wenn } (n_{Parolen}[i] = \text{Parolen}[i]), \text{ dann } (Stimmenanteil))$$



## 4.5 Design der App

Das Design der App ist ein Arbeitsschritt, welcher einfach zu übersehen ist und in den vorhergehenden Teilen auch kaum thematisiert wurde. Allerdings sollte er doch angesprochen werden, da relativ viel Arbeit in das Aussehen der App floss. Einige Grundsätze legte ich schon vor dem Anfang des Programmierens fest. Jede Funktionsseite würde in zwei Teile geteilt sein. Links eine Seitenleiste, auf welcher Einstellungen gemacht werden können, rechts eine grössere Fläche, um jeweils die Resultate anzeigen zu können. Auch eine Farbpalette legte ich fest. Für Elemente und Hintergrund zwei verschiedene dunkle Grautöne, sowie einen weiteren Grauton zur Hervorhebung einer ausgewählten Option. Ausserdem bestimmte ich einen Rotton und einen Blauton. Diese wurden einerseits für die Republikaner und Demokraten genutzt, andererseits aber auch für die Nein- und Ja-Seiten bei Schweizer Abstimmungen. Auch die Schriftart wollte ich einheitlich halten. Für die meisten Elemente konnte ich auch die gewählte Monospace-Schriftart *Menlo* nutzen.

Dies ist auch schon ein erstes Problem mit Xamarin. Einige integrierte Steuerelemente lassen sich nur bedingt anpassen. Dies liegt unter anderem daran, dass Xamarin auf mehreren Plattformen laufen muss, wobei nicht alle die gleichen Funktionalitäten haben. Dies führt zu einigen Problemen. So kann beispielsweise das Button-Steuerelement höchstens eine einzige Zeile Text anzeigen. Dies kann durch sogenannte *Custom Renderers* gelöst werden. Diese überschreiben die Funktionen, welche normalerweise genutzt werden, um die Steuerelemente darzustellen. Solche *Custom Renderers* waren auch notwendig, nachdem das Update auf iOS 15 gemacht wurde. Dieses veränderte nämlich die Eigenschaften einiger Steuerelemente, welche dann nicht mehr richtig funktionierten.

Neben diesen eigenen Darstellungsmethoden musste ich teilweise auch eigene Elemente implementieren. So wird zum Beispiel der Graph in der Modell-Funktion durch eine eigene Klasse erstellt. Auch simplere Elemente, wie zum Beispiel die Dreifachauswahlknöpfe, welche sowohl für den Parolenbetrachter als auch das Modell benutzt wurden, mussten selbst erstellt werden und sind separate Klassen. Auch die Listen-Auswahl-Elemente für das Modell habe ich selbst erstellt. Zwar gibt es ein Standard-Listen-Auswahl-Element, welches ich auch für den Wahldatenbetrachter benutzt habe, allerdings ist es in der Bedienung umständlich und lässt sich beispielsweise betreffend Schriftart nicht anpassen. Auch zum Design zählen die visuellen Ressourcen, die in der App hinterlegt sind. Einerseits sind dies die Karten, welche im Vektorformat abgelegt sind, andererseits aber auch die Partei- und Organisationslogos, welche in der App jeweils neben den Namen der betreffenden Organisationen erscheinen. Diese mussten alle einzeln gesucht und gegebenenfalls angepasst werden. Nicht zuletzt aufgrund der Mühsamkeit dieses Prozesses erstellte ich sowohl das App-Icon als auch die Icons für die einzelnen Teilbereiche in einem Vektorzeichnungsprogramm selbst.

## 5. Diskussion

### 5.1 Resultate des Modells

Um die Vorhersagen des Modells überprüfen zu können, muss eine gewisse Basis etabliert werden; es muss also eine Alternative etabliert werden, mit welcher das Modell verglichen werden kann. Hierfür wurden mehrere alternative Methoden ausprobiert:

Nummer	Name	Beschreibung
1.	Einfachste Annahme	Das einfachste mögliche Modell. Es wird in allen Fällen die einfachste Annahme getroffen, als Vorhersage wird das durchschnittliche Resultat über alle Vorlagen angenommen. Über alle Volksabstimmungen ist das durchschnittliche Resultat 50% Ja-Anteil.
2.	Rechtsform-Vorhersage	Ein häufiger Indikator ist die Rechtsform einer Vorlage. Initiativen werden meistens abgelehnt, Referenden meist angenommen. Deswegen wird als Vorhersage für eine Initiative die durchschnittliche Zustimmung einer Initiative (ca. 35%) und für das Referendum die durchschnittliche Zustimmung bei einem Referendum (ca. 51%) angegeben.
3.	Parolen-Vorhersage	Diese Methode beruht nur auf den Partei-Parolen. Die Stimmen-Prozentsätze der befürwortenden Parteien werden addiert und bilden die Vorhersage. Die Annahme ist, dass alle Abstimmenden genau den Parolen der Partei folgen, für welche sie bei den Nationalratswahlen gestimmt haben.
4.	Vollständiges Modell	Das vollständige in 4.4 <i>Erstellung des Modells</i> erklärte Modell, welches das Jahr, die Rechtsform, den Politikbereich und die Parolen als Auswahlkriterien beinhaltet und Annahmen über deren jeweilige Gewichtung bei Bestimmung der Ähnlichkeit macht.

Obwohl basierend auf den Parolen und der Rechtsform jeweils ein simples Modell entwickelt werden konnte, wäre dies für das Jahr und den Politikbereich wenig sinnvoll. Ein Modell, welches nur auf dem Jahr basiert, würde entweder eine Aussage machen, die sich nahe an 50% befindet oder nur die letzten wenigen Abstimmungen reflektieren, welche aber keine Aussagekraft zu zukünftigen Abstimmungen besitzen. Eine Einordnung nach Politikbereich scheitert einerseits daran, dass es jeweils wenige Daten zu einem bestimmten Bereich gibt und der Bereich noch nichts über die Richtung des Inhalts aussagen kann. Nur weil eine Vorlage zur Finanzpolitik gehört, weiss man noch nicht, ob sie beispielsweise höhere Steuern befürwortet oder ablehnt. Rechtsform und Parolen korrelieren mit den Resultaten, Jahr und Politikbereich tun dies nicht.

Die vier Methoden wurden stichprobenmässig ausgewertet. Evaluiert wurden die Methoden jeweils für die Zeitperioden ab 1980, 2000 und 2020. In den jeweiligen Jahren wurden ab der ersten Abstimmung des Jahres jeweils die 20 darauffolgenden Abstimmungen ausgewertet. Diese ersten Abstimmungen waren jeweils auch die ersten Abstimmungen nach der letzten Parlamentswahl (1979, 1999 und 2019).

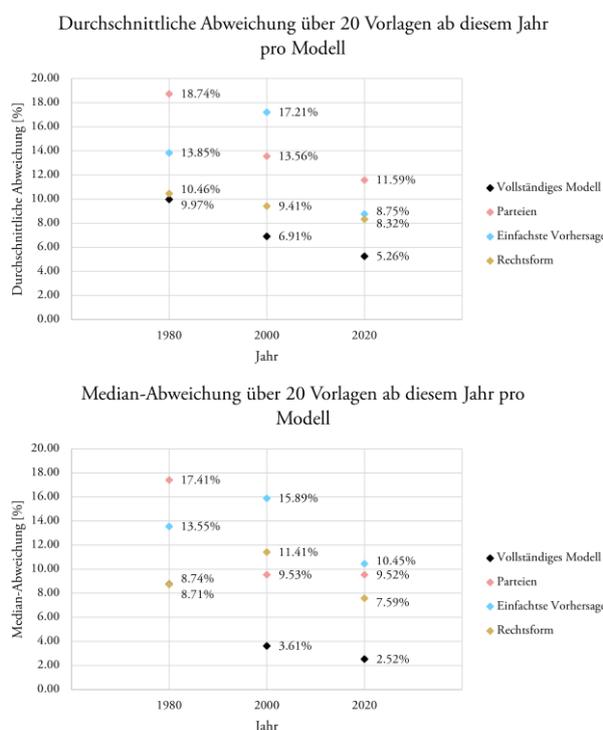


Abb. 8: Die beiden Graphen zeigen oben die mit den vier Methoden erzielten durchschnittlichen Abweichungen und unten die damit erzielten Median-Abweichungen. Jeder Datenpunkt entsteht aus einem Durchschnitt von zwanzig ab Beginn des Jahres in der x-Achse stattgefundenen Abstimmungen. So haben die mit dem vollständigen Modell errechneten Vorhersagen für zwanzig ab 2020 stattgefundenen Abstimmungen durchschnittlich eine Abweichung von 5.26% vom echten Resultat. Bemerkenswert ist, dass die Vorhersagen des vollständigen Modells, der Rechtsform- und der Parteien-basierten Vorhersage mit den Jahren alle besser werden.

Genau 20 Abstimmungen wurden aus dem Grund gewählt, dass seit 2019 im Datensatz so viele Abstimmungen stattgefunden haben. Durch eine grössere Anzahl würde für die Evaluation ab 1980 ausserdem bereits viele Vorlagen miteinbezogen, welche erst nach den nächsten Parlamentswahlen 1983 stattgefunden haben. Die grossen zeitlichen Abstände zwischen den drei Stichproben wurde gewählt, damit sich die Zeitpunkte möglichst wenig gegenseitig beeinflussen. Die beiden Graphen in Abbildung 8 zeigen die mit den jeweiligen Modellen erzielten durchschnittlichen und Median-Abweichungen von den echten Resultaten über die jeweils analysierten 20 Vorlagen. Dabei erzielt das vollständige Modell die tiefsten Abweichungen. Bemerkenswert ist ausserdem, dass das vollständige Modell, das Parolen-basierte Modell und das Rechtsform-basierte Modell über die Jahre allesamt kleinere Abweichungen erzielten.

Hierfür sind viele Gründe möglich. Einerseits zeigt sich, dass beispielsweise das vollständige Modell mit mehr Daten bessere Resultate erzielt, diese Entwicklung zeigt sich aber mit deutlich weniger zugrundeliegenden Daten. Dies ist in Abbildung 9 ersichtlich. Das Modell macht mit 10 vorherigen Vorlagen als Input bessere Vorhersagen, als wenn diese Vorhersage auf nur eine vorherige Abstimmung basiert. Der Sprung von 100 auf 200 Vorlagen als Input ist aber deutlich kleiner, die Vorhersagen werden nur minimal besser. Als Grund für die über die Jahre besseren Vorhersagen ist diese Erklärung also nicht sinnvoll. Anscheinend folgen die Bürger eher den Parolen ihrer Partei und hören vermutlich auch eher auf die Behörden. Auch die Zeit von sehr klaren Abstimmungen und einer stark parteiischen Presse sind vorbei.

Konfiguriert wurden die Gewichtungen des vollständigen Modells anhand von 200 Abstimmungen, welche vor dem Evaluationszeitraum stattfanden. Für die 20 Abstimmungen ab 2020 waren dies die Abstimmungen von 1997 bis 2019. Werden weniger viele frühere Abstimmungen zur Optimierung der Gewichtungen benutzt, ändern sich diese nur wenig (Exakte Daten finden sich im Excel-Dokument auf Github). Es fällt vor allem auf, dass wenn weniger Daten zur Konfiguration der Gewichtung benutzt werden das Gewicht für das Jahr sehr klein wird. Dies ist sinnvoll: Die kleinere Datenmenge erstreckt sich über eine deutlich kleinere Zeitspanne, ältere Abstimmungen müssen also nicht aktiv «ignoriert» werden. Das Gewicht der Partei-Parolen wird grösser, wenn die Menge der zur Konfiguration der Gewichte benutzten Abstimmungen kleiner wird. Dies könnte daran liegen, dass die Partei-Prozente in der analysierten Datenmenge durch die kleinere Zeitspanne weniger weit streuen und damit aussagekräftiger werden.

Das vollständige Modell liegt mit seinen Vorhersagen in gut 80% der Fälle richtig (77% in 2021, 78% in 2020, 82% von 1997 bis 2019). Beim Vergleich der durchschnittlichen Abweichung und der Median-Abweichung fällt auf, dass die Median-Abweichung meist kleiner wie die durchschnittliche Abweichung ist. Dies liegt vor allem daran, dass einige sehr grosse Abweichungen den Mittelwert in die Höhe trieben. In der

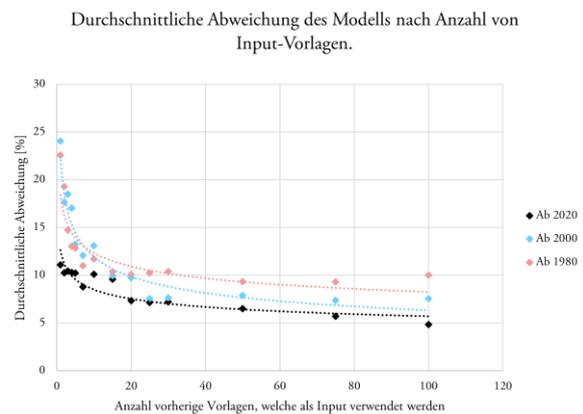


Abb. 9: Dieser Graph bildet für das vollständige Vorhersage-Modell die durchschnittliche Abweichung der Vorhersage vom echten Resultat abhängig von der Anzahl zur Vorhersage benutztem Vorlagen als Input ab, wobei die Analyse jeweils für die 20 ersten Abstimmungen in den Jahren 1980, 2000 und 2020 durchgeführt wurde. Wenig überraschend ist, dass die durchschnittliche Vorhersage besser wird, wenn mehr Daten verfügbar sind. Spannend ist aber, wie schnell sich die Vorhersagen verbessern. Umso mehr Vorlagen schon zum Machen der Vorlage verwendet werden, umso weniger bringt es, mehr Daten hinzuzufügen. Um die für die Vorhersage zur Verfügung stehenden Vorlagen zu bestimmen, wurden die Vorlagen benutzt, welche direkt vor der zu vorhersagenden Abstimmung stattfanden. Es wurde also rückwärts von der zu vorhersagenden Abstimmung gezählt. In der normalen Version des Modells sind alle vorherigen Abstimmungen für eine Vorhersage verfügbar.

Kontrollmenge von zwanzig Abstimmungen ab 2020 sind die drei klarsten Beispiele das E-ID-Gesetz, das CO<sub>2</sub>-Gesetz und die Pflegeinitiative. Für alle diese Fälle gab es «unerwartete» Elemente. Gegen das E-ID-Gesetz wurde von den Gegnern überraschend effektiv politisiert, das CO<sub>2</sub>-Gesetz scheiterte vor allem aufgrund von zwei gleichzeitig stattfindenden Abstimmungen zu Initiative in einem ähnlichen Politikbereich. Die Ablehnung gegen diese beiden Initiativen brachte auch das Gesetz zum Scheitern. Die Pflegeinitiative wurde mit überraschend hohen Anteilen angenommen. Dies lag vor allem an der vergleichsweise temporär sehr hohen Prominenz des Pflegepersonals wegen der Covid-19-Pandemie.

Das stärkste Element im Modell ist aber, dass es zwar eine Vorhersage abgibt, allerdings sehr viele andere Möglichkeiten vorhersieht. So ist beispielsweise die Vorhersage für das Gesetz «Ehe für alle» sehr nahe am echten Resultat, die Abweichung beträgt ca. 1.3%. Trotzdem errechnet das Modell in nur circa 73% der Fälle, dass die Vorlage angenommen wird. Es schreibt also dem Ereignis «Vorlage wird angenommen» in diesem Fall eine Chance von ca. 73% zu. Die Erwartung wäre nun, dass in gut 70% der Fälle, in welchen das Modell annimmt, dass die Vorlage mit ca. 70%-Chance angenommen wird, diese Vorlage auch wirklich angenommen wird. Es zeigt sich, dass das Modell tendenziell zu vorsichtig ist, dem wahrscheinlichsten Resultat also eine zu kleine und weniger wahrscheinlichen Resultaten eine zu hohe Wahrscheinlichkeit zu-teilt.

## 5.2 Allgemeines Fazit

Ich hatte am Anfang der Arbeit mehrere Hypothesen aufgestellt, welche sich alle mehr oder weniger gut zeigen liessen. Allerdings hatte ich sie auch relativ vage formuliert. Die erste Hypothese, dass in den Vereinigten Staaten deutlich mehr verschiedene Methoden zur Abstimmungsvorhersage existieren, konnte ich meiner Meinung nach im ersten Teil der Arbeit zufriedenstellend zeigen. In der Schweiz ist die weitaus populärste Methode, eine Abstimmung vorherzusagen, sich auf die Umfragen zu verlassen. In Amerika gibt es einerseits Modelle, welche aus den Umfragedaten nochmals deutlich mehr Informationen herausholen können, andererseits aber auch Vorhersagemethoden, welche von Umfragewerten unabhängig funktionieren.

Meine zweite Hypothese knüpfte gleich daran an. Diese war, dass sich diese Modelle auf die Schweiz übertragen lassen sollten. Diese Hypothese erfüllte sich nur bedingt. Das Modell, welches ich schlussendlich entwickelt habe, ist nicht analog zu einem einzelnen System in den USA. Es nutzt sowohl grundlegende Faktoren, wie das *13-Keys*-Modell, aber auch computergestützte Simulationen, wie das *fivethirtyeight*-Modell. Es war mir klar, dass gewisse Abänderungen nötig wären, trotzdem waren diese grösser als erwartet. Allerdings stellten sich die Informationen, welche in der Analyse der amerikanischen Systeme sichtbar wurden, als sehr wichtig für die Entwicklung des Modells heraus. Ohne meine Recherchen wäre ich kaum in der Lage gewesen, die Funktionsweise des Vorhersagemodells zu erarbeiten und umzusetzen.

Die dritte Hypothese war, dass ein Modell in der Lage sein sollte, Schweizer Volksabstimmungen zufriedenstellend genau vorherzusagen und eine ähnliche Abweichung wie Umfragen hat. Diese Hypothese kann als erfüllt bezeichnet werden, auch wenn die als Kontrolle benutzten Abstimmungen (drei Mal jeweils zwanzig) noch nicht genügend viele sind, um ausschliessen zu können, dass das Modell einfach in vielen Fällen Glück hatte. Auch hat das Modell noch klare Fehler, allerdings sind diese auch nachvollziehbar genug, um

vermuten zu können, dass sie behoben werden können. Eine weitere Optimierung war mit dem Zeitplan allerdings nicht vereinbar. Auch gibt es viele Dinge, die ich gerne noch weiter untersucht hätte, welche aber aus verschiedenen Gründen nicht umgesetzt werden konnten. So hätte ich gerne auch ein Modell programmiert, welches mit Umfragedaten rechnet, und eine Vorhersage des Meinungsbildungsprozesses macht. Bei diesem Beispiel fehlten mir aber die entsprechenden historischen Umfragedaten.

Meine Fragestellung nach der Möglichkeit der Übertragung amerikanischer Modelle kann ich betreffend des «ob» mit ja beantworten, während das «wie» in der vorhergehenden Arbeit thematisiert wird. Ein weiterentwickeltes Modell könnte als Zusatz zu Umfragen dienen. Allerdings kann es immer nur einen unterstützenden Charakter haben. Die Natur einer Umfrage erlaubt es ihr, deutlich mehr Informationen einzufangen. Dabei handelt es sich um Einflüsse, die nur schwer von einem Modell betrachtet und erfasst werden können, insofern sie das erste Mal auftreten. Allerdings ist nur schon die für das Modell grundlegende Funktion, ähnliche Abstimmungen zu finden, von gewissem Wert, da diese Abstimmungen auch unabhängig vom Modell als Anhaltspunkte dienen können. Das Wichtigste ist, dass die Resultate von Modellen und Umfragen nie einfach so als Wahrheit akzeptiert werden dürfen, sondern die in der Vorhersagetechnik grundlegenden Unsicherheiten immer zentral erwähnt werden sollten. Allerdings ist es auch nicht so, dass es sich um ein Rätselraten handelt. Mein Modell, aber auch die gfs.bern-Umfragen, haben einige erstaunlich genaue Resultate produziert. Die Betonung der Unsicherheit ist für die Fälle wichtig, in denen sie es nicht tun.

## 6. Schlusswort

### 6.1 Zusammenfassung

Diese Arbeit setzt sich mit Wahlvorhersagen auseinander, spezifisch der Übertragung von in Amerika benutzten Modellen auf Schweizer Abstimmungen. Hierfür habe ich erst die Entwicklung der Wahlvorhersagen in den Vereinigten Staaten untersucht und im ersten Teil dieser Arbeit dokumentiert.

Dieser erste Teil war vor allem eine historische Recherche. In zwei Unterkapiteln erörtere ich die Entstehung von Umfragen vor Wahlen, welche ihren Ursprung als lokale Befragung in verschiedensten Ortschaften hatten. Später wurden sie landesweit durchgeführt, allerdings mit heutzutage eindeutig veralteten Methoden. Ein Magazin aus New York wurde für seine Umfragen vor Präsidentschaftswahlen bekannt, ihre Methodik bedingte aber, dass es nicht möglich war, zu bestimmen, welche Personen und Personengruppen an der Umfrage mitgemacht hatten. Dies führte 1936 zu einer katastrophale Fehlvorhersage. Abgelöst wurde dieses Magazin durch professionelle Marktforschungsinstitute, welche sich auf wissenschaftliche Methoden berufen. Auch diese Methoden hatten noch einige Fehler und führten wiederum zu Fehlvorhersagen. Die Methoden verbessern sich weiterhin. In einem weiteren Teilkapitel thematisiere ich Wahlvorhersagen, welche ohne Umfragedaten auskommen. Hierfür ist die Betrachtung historischer Begebenheiten grundlegend. Das bekannteste Modell dieser Art zeigt nur den Gewinner der Wahl an, nicht aber ein erwartetes Resultat. Es kann gezeigt werden, dass sich dieses Modell aber auch leicht umformen lässt, um eine bessere Vorhersage zu erreichen. In einem weiteren Unterkapitel thematisiere ich eine neue Entwicklung in der Wahlvorhersage, welche auf Computersimulationen basiert und als Basisdaten vor allem Umfragen nutzt, aber auch von anderen Indikatoren Gebrauch macht. Diese Methode hat einerseits bereits sehr genaue Vorhersagen produziert, andererseits vermag sie aber auch die Unsicherheiten, die in Vorhersagen grundlegend sind, zu beinhalten und zu kommunizieren.

Zweiter Hauptbestandteil der Arbeit ist die Entwicklung einer App. Die Entwicklung verarbeitete ich in zwei Kapiteln. Das erste zeigt das Konzept der jeweiligen App-Bestandteile, im zweiten Kapitel wird die Entwicklung und Funktionsweise dokumentiert und erklärt. Die App ist in drei Funktionsbereiche unterteilt. Einerseits sind dies zwei kleinere Bestandteile, welche vor allem zur Visualisierung von Daten nützlich sind. Die erste Funktion der App ist ein Datenbetrachter, in welchem die für den ersten Teil der Arbeit relevanten Abstimmungsdaten ersichtlich sind. Ausserdem sind hier kleine Analysefunktionen eingebaut, welche rudimentäre Veränderungen an und Vergleiche zwischen den Daten zulassen. Der zweite Teil ist ein Tool zur Analyse von Parteiparolen. Dieser Teil diente vor allem als Grundlage für den dritten Teil und hat eigenständig nur wenig Aussagekraft. Der dritte Funktionsbereich der App ist das Kernstück dieser Arbeit. Es handelt sich um ein Vorhersagemodell, welches auf den Daten früherer Volksabstimmungen basiert und so Vorhersagen über eine neue Abstimmung machen kann. Bei der Analyse der Vorhersagen des Modells für die zum Abgabzeitpunkt sieben neusten Abstimmungen kommt eine beeindruckend kleine Abweichung heraus. Allerdings gibt es auch noch einige eindeutige Fehler im Modell, welche weitere Anpassungen benötigen. Diese sind aber behebbbar. Es sollte also möglich sein, ein weiter optimiertes derartiges Modell für Vorhersagen zu nutzen.

## 6.2 Persönliche Erkenntnisse

Im Rahmen meiner Arbeit habe ich viele Dinge gelernt. Es war eine langwierige, aufwändige, aber auch sehr interessante Arbeit. Ich konnte Fähigkeiten aus diversen Fachgebieten nutzen, um die Fragestellung und Hypothesen zu erarbeiten. Natürlich involvierte dies sowohl Informatik als auch Geschichte und Staatskunde, allerdings sind auch andere Fähigkeiten nicht zu vernachlässigen. Beispielsweise fällt auf, dass fast mein gesamtes Quellenverzeichnis in englischer Sprache ist. Dies liegt daran, dass ich mich an amerikanischen Vorbildern orientiert habe und deswegen Quellen meist nur auf Englisch zur Verfügung standen.

Die historische Recherche war oftmals ein wenig aufwändig, aber nicht weniger interessant. Es war spannend, die verschiedenen Primärquellen zu finden, auch wenn die verschiedenen Online-Archive oftmals nur schwer bedienbar sind. Allerdings habe ich dadurch auch gelernt, wie ich in Zukunft schneller auf diese Ressourcen zugreifen kann. Ausserdem war mir nicht wirklich klar, welche Einblicke in die Geschichte Zeitungsartikel geben können. Ich hatte schon zuvor Originalquellen zu wichtigen Ereignissen gelesen, und fand diese bereits sehr interessant. Gerade aber frühe Lokalzeitungen geben einen nochmals tieferen Einblick in das Leben in der damaligen Zeit. Die erste Wahlvorhersage war in einer Zeitung, welche relativ viele Inserate abdruckte. Beispielsweise ist dort davon zu lesen, dass einem Herrn seine Peitsche abhandengekommen ist und der Finder sie doch möglichst schnell retournieren solle. Auch ein Leserbrief, in welchem sich jemand in einer faszinierend eloquenten und pompösen Sprache über die Benutzung von Schimpfwörtern aufgeregt, ist zu finden. Allerdings sind auch Inserate abgedruckt, welche heute ein wenig gar makaber wirken. So wurden 125 Dollar als Belohnung ausgeschrieben, sollten einem Besitzer seine entlaufenen Sklaven zurückgebracht werden. Primärquellen ermöglichen spannende Einblicke in die Umgangsformen und Selbstverständlichkeiten anderer Zeiten. Diese Einblicke waren zwar nicht zwingend für die Arbeit relevant, haben meiner Meinung nach aber meine Fähigkeit zum Umdenken in andere Zeiten verbessert.

Auch meine Programmierfähigkeiten haben sich definitiv im Laufe der Arbeit verbessert, und ich habe wertvolle Einblicke erhalten, was alles in die Produktion und Publikation von Software involviert ist. Gerade der Apple App Store hat teilweise recht strenge Vorschriften, nicht zuletzt auch betreffend Design, was eigentlich nicht meine Kernkompetenz ist. Allerdings hat mir gerade auch dies Spass gemacht. Einerseits war es eine willkommene Abwechslung, andererseits habe ich an der Kantonschule auch immer die Breite des vermittelten Wissens wertgeschätzt. Deswegen war es schön, Fähigkeiten aus den verschiedensten Fachbereichen nutzen zu können.

Allerdings war nicht alles an der Arbeit so schön, wie ich es hier ausmale. Meine Zeitplanung ist oftmals mehr wie nur suboptimal. Deswegen könnte gerade der Text-Teil meiner Arbeit meiner Meinung nach in einigen Teilen ergänzt und verbessert werden. Die Programmierung und Recherche habe ich deutlich weniger vernachlässigt wie den Schreibprozess. Deswegen bin ich dort auch deutlich zufriedener mit dem Endresultat. Allgemein verbuche ich die Erfahrung definitiv als positiv. Auch, dass mein Modell überraschend gut funktioniert hat, werte ich als Erfolg. Den nächsten Abstimmungen blicke ich aber trotzdem recht unsicher entgegen.

## 7. Benutzeranleitung

Die entwickelte App kann über den App Store auf jedes iPad heruntergeladen werden, welches Softwareversion iPad-OS 12.0 oder höher hat. Allerdings wurde die App nur auf den Versionen 14 und 15 getestet, weswegen nur für diese Versionen ein Funktionieren garantiert werden kann. Ausserdem sind auf dem Gerät gut 50 Megabyte freier Speicherplatz nötig. Nach dem Download der App benötigt diese kein Internet, alle Daten sind hinterlegt und abgespeichert. Die App ist unter dem Namen *Election Predict* im App Store auffindbar. Alternativ kann auch der QR-Code in Abbildung 8 gescannt werden.

Es ist auch möglich, den Code von Github herunterzuladen, und das Projekt selbst zu kompilieren. Hierfür sind funktionierende Installationen von Visual Studio und X-Code notwendig. Auch nötig sind die entsprechenden Apple-Entwickler-Lizenzen. Den gesamten Prozess hier auszuführen würde den Rahmen dieser Anleitung sprengen. Sollte jemand erwägen, das Projekt selbst zu kompilieren, kann dieser Person auch vertraut werden, die Konfiguration selbst aufzusetzen.

Ist die App geöffnet, stehen die drei Funktionsbereiche mit der unteren Navigationsleiste zur Verfügung. Um Daten zu betrachten, kann der Tab «Wahldaten» angewählt werden. Im Menu kann dann ein Land, dann eine Publikation, oder im Falle der Schweiz ein Abstimmungsdatum, ausgewählt werden. Schliesslich kann der genaue Datenpunkt ausgewählt werden. Bei entsprechenden Daten erscheinen zwei weitere Optionen. Unter «Vergleichen» kann ein weiterer Datenpunkt ausgewählt werden, unter «Modifizieren» können die Resultate der Datenpunkte mit einem Schieberegler verändert werden. Mit dem Anwählen des «Berechnen»-Knopfes werden die Daten auf der rechten Bildschirmseite angezeigt.

Mit der Auswahl des Tabs «Partei-Parolen» gelangt man zu der Parolen-Analyse. Linkerhand befinden sich alle Organisationen, welche analysiert werden können. Werden die jeweiligen Schaltflächen angewählt, erscheint auf der rechten Bildschirmseite eine Schaltfläche mit der Organisation. Mit dem Antippen des Kreuzes oben rechts auf der Schaltfläche kann diese wieder entfernt werden. Mit dem sich auf der Schaltfläche befindlichen Selektor können die Parolen «Ja», «Nein» und «Neutral» ausgewählt werden. Die Schaltfläche neutral zeigt an, dass eine Parole beispielsweise die Stimmfreigabe bestimmte. Um eine Organisation aus der Berechnung zu entfernen, muss das Kreuz oben rechts angewählt werden. Sobald die ersten Parolenauswahlen getroffen wurden, erscheinen oben auf der rechten Bildschirmseite die relevanten Kennzahlen.

Mit der Auswahl des Tabs «Modell» gelangt man zum Vorhersagemodell. Unter dem Bereich «Jahr», kann das Jahr der zu vorhersagenden Abstimmung gewählt werden. Unter «Rechtsform» kann angewählt werden, um welche Art der Abstimmung es sich handelt. Unter «Politikbereich» muss erst das allgemeine Thema ausgewählt werden. Ist dies getan, erscheint eine weitere Schaltfläche mit untergeordneten Bereichen. Ist dort ein Bereich ausgewählt, erscheint meist eine weitere Schaltfläche mit weiter verfeinerten Auswahlmöglichkeiten. Im Bereich «Parolen» können dann für die wichtigsten Parteien die Parolen ausgewählt werden.

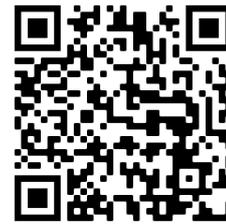


Abb. 10: QR-Code mit Link zum App Store.

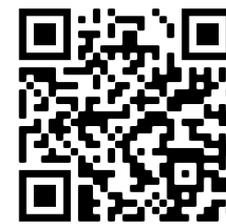


Abb. 11: QR-Code mit Link zum Github Repository.

Die Option «Keine Auswahl» entfernt sie hier aus der Berechnung und steht nicht explizit für eine Stimmfreigabe seitens der Partei. Allerdings sollte sie in diesem Fall trotzdem genutzt werden. Der Berechnungsprozess kann mit dem Knopf «Berechnen» gestartet werden, die Resultate erscheinen rechterhand.

Eine ausführliche Benutzeranleitung ist zusammen mit der Dokumentation auf Github verfügbar.

## 8. Quellen

### 8.1 Primärliteratur und Datenquellen

- American Watchman, Ausgabe 16. Juli 1824, Wilmington, Delaware  
 The Evening Star, Ausgabe 4. November 1948, Washington DC  
 The Gallup Poll: Public Opinion 1935 – 1971 Vol. 1 – 3, 1972, New York  
 The Literary Digest, New York City:  
     Ausgabe 16. September 1916  
     Ausgabe 31. September 1936  
 The New York Times, New York City:  
     Ausgabe 4. Oktober 1924  
     Ausgabe 19. Juli 1936  
     Ausgabe 1. November 1936  
     Ausgabe 4. November 1948  
     Ausgabe 16. November 1980  
 The San Francisco Call, San Francisco:  
     Ausgabe 6. November 1912  
 Time Magazine, New York City:  
     Ausgabe 23. Mai 1938  
     Ausgabe 3. Mai 1948

### 8.2 Sekundärliteratur

- Adams, William C.: *Election Night News and Voter Turnout*, 2005, ISBN 9781588263810  
 Atanasov, Pavel et al: *Distilling the Wisdom of Crowds: Prediction Markets vs. Prediction Polls*, 2016,  
<https://doi.org/10.1287/mnsc.2015.2374>  
 Campbell, W. Joseph: *Lost in a Gallup*, 2020, ISBN 9780520300965  
 Chambers, William Nisbet: *Party Development and the American Mainstream* in *The American Party System: Stages of Political Development*, 1975, ISBN 9780195019179  
 Erikson, Robert S. et al: *Likely (and unlikely) voters and the assessment of campaign dynamics* in *Public Opinion Quarterly*, 2004, <https://doi.org/10.1093/poq/nfh041>  
 Erikson, Robert S. & Wlezien, Christopher: *Are Political Markets Really Superior to Polls as Election Predictors?* In *Public Opinion Quarterly*, 2008, <https://doi.org/10.1093/poq/nfn010>  
 Lichtman, Allan J.: *Predicting the next President*, 2020, ISBN 9781538148655  
 Longchamp, Claude: *Prädisposition bei Schweizer Volksabstimmungen* in *Direkte Demokratie: Forschung und Perspektiven*, 2002, ISBN 9783531138527  
 Palfrey, Thomas R.: *A Mathematical Proof of Duverger's Law* in *Models of Strategic Choice in Politics*, 1989, <https://doi.org/10.5860/choice.28-0581>  
 Perry, Paul: *Election Survey Procedures of the Gallup Poll* in *Public Opinion Quarterly*, 1960, <https://doi.org/10.1086/266966>  
 Rigdon, Steven E. et al: *A Bayesian Prediction Model for the U.S. Presidential Election* in *American Politics Research*, 2009, <https://doi.org/10.1177/1532673x08330670>  
 Senti, Martin: *Wetterleuchten der Demoskopie* in *Wahl-Probleme der Demokratie*, 2012, ISBN 9783725564927  
 Tankard, James W.: *Public Opinion Polling by Newspapers in the Presidential Election Campaign of 1824* in *Journalism Quarterly*, 1972, <https://doi.org/10.1177/107769907204900219>  
 Wilks, Samuel S. et al: *Report on the Analysis of Pre-Election Polls and Forecasts* in *Public Opinion Quarterly*, 1949, <https://doi.org/10.1086/266004>

### 8.3 Internetartikel

- Enten, Harry: *Gallup Gave Up. Here's Why That Sucks*, 7. Oktober 2015, <https://fivethirtyeight.com/features/gallup-gave-up-heres-why-that-sucks/>, abgerufen am 7. August 2021  
 Katz, Josh: *Who Will Be President?*, 8. November 2016, <https://www.nytimes.com/interactive/2016/upshot/presidential-polls-forecast.html>, abgerufen am 11. September 2021  
 Kovic, Marko: *Longchamps SRG-Umfragen im 10-Jahres-Check*, 9. Dezember 2014, <https://blog.tagesanzeiger.ch/daten-blog/index.php/6760/longchamps-srg-umfragen-im-10-jahres-check>, abgerufen am 10. Oktober 2021  
 Silver, Nate: *How FiveThirtyEight's 2020 Presidential Forecast Works – And What's Different Because of Covid-19*, 12. August 2020, <https://fivethirtyeight.com/features/how-fivethirtyeights-2020-presidential-forecast-works-and-whats-different-because-of-covid-19>, abgerufen am 11. September 2021

Thomas, Andrew: *538s Uncertainty Estimates Are As Good As They Get*, 7. November 2012, <https://web.archive.org/web/20121115083320/http://www.acthomas.ca/comment/2012/11/538s-uncertainty-estimates-are-as-good-as-they-get.html>, abgerufen über archive.org am 7. Oktober 2021

#### 8.4 Internetressourcen

App im App-Store: <https://apps.apple.com/ch/app/election-predict/id1564505507>

Github-Repository mit dem Quellcode: <https://github.com/AHU03/ElectionPredict>

Math.NET Numerics-Paket, für die Entwicklung der App genutzt: <https://numerics.mathdotnet.com/>

Roper iPoll, für Gallup-Daten in der App (kostenpflichtig): <https://ropercenter.cornell.edu/ipoll/>

SkiaSharp-Paket, für die Entwicklung der App genutzt: <https://docs.microsoft.com/en-us/xamarin/xamarin-forms/user-interface/graphics/skiasharp/>

Swissvotes-Datensatz: <https://swissvotes.ch/page/dataset>

#### 8.5 Abbildungsverzeichnis

Abbildung 1: American Watchman: <https://chroniclingamerica.loc.gov/lccn/sn82014894/1824-07-16/ed-1/seq-3/>, Bild digital heller gemacht

Abbildung 2: Literary Digest: <https://archive.org/details/literarydigest53newy/page/659/mode/2up>, Bild digital heller gemacht

Abbildung 3: Quelle: Eigene Darstellung mit Daten aus *The Gallup Poll* und Roper iPoll für die Gallup-Daten, sowie Daten aus der Encyclopaedia Britannica für die echten Resultate. Fehlgeschlagene Vorhersagen nach Campbell: *Lost in a Gallup*

Abbildung 4: Daten aus Allan J. Lichtmans *Predicting the Next President*, sowie der Encyclopaedia Britannica für die Stimmen-Prozentsätze. Eigene Darstellung, eigene Berechnungen.

Abbildungen 5, 6 und 7: Screenshots aus der App, von einem iPad 5, iPad-OS Version 14. Eigene Bilder.

Abbildungen 8 und 9: Graphen mit Daten aus den Modellen, erstellt mit Excel. Eigene Bilder

Abbildungen 10 und 11: QR-Codes, erstellt mit <https://qr-code-generator.com>