

'Nearest Neighbour'

Prediktor (NNP)

als

JAVA<sup>™</sup> - Applet

von

Dr. rer. nat. Hans Uhlig, Wissenschaftsautor

## Copyright Hinweis

Der Text und die Abbildungen dieses Beitrages sowie die zugehörige Software unterliegen dem Urheberrechtsschutz. Wer diese Produkte erwirbt, darf sie für den eigenen Gebrauch nutzen. Kopien oder Wiedergaben in anderer Form ob vollständig oder nur teilweise, bedürfen meiner schriftlichen Zustimmung.

Hans Uhlig, Oktober 2010

## Einführung

Ein Nearest Neighbour Prediktor, der mit einem Kalkulationsblatt erstellt wird, hat den großen Vorteil, dass alle Berechnungen und ihre Ergebnisse sehr transparent und jederzeit nachvollziehbar sind. Zum Kennenlernen der Methode ist diese Art der Umsetzung daher ideal. Wenn man jedoch mehrere Varianten ausprobieren möchte, ist der dafür nötige Zeitaufwand einfach viel zu groß. Mit dem beigefügten JAVA™-Applet können Sie gleichzeitig in sehr viel kürzerer Zeit mehrere Varianten des Prediktors ausprobieren. Es werden beispielsweise immer gleichzeitig sechs Vorhersagevarianten angezeigt. Es sind die Vorhersagen mit 3,4 und 5 nächsten Nachbarn und zwar in der oberen Reihe, siehe Bildschirmausdruck, mit einheitlicher Gewichtung und in der unteren mit ihrem jeweiligen Abstand (Distanz) zum Referenzmuster gewichtet.

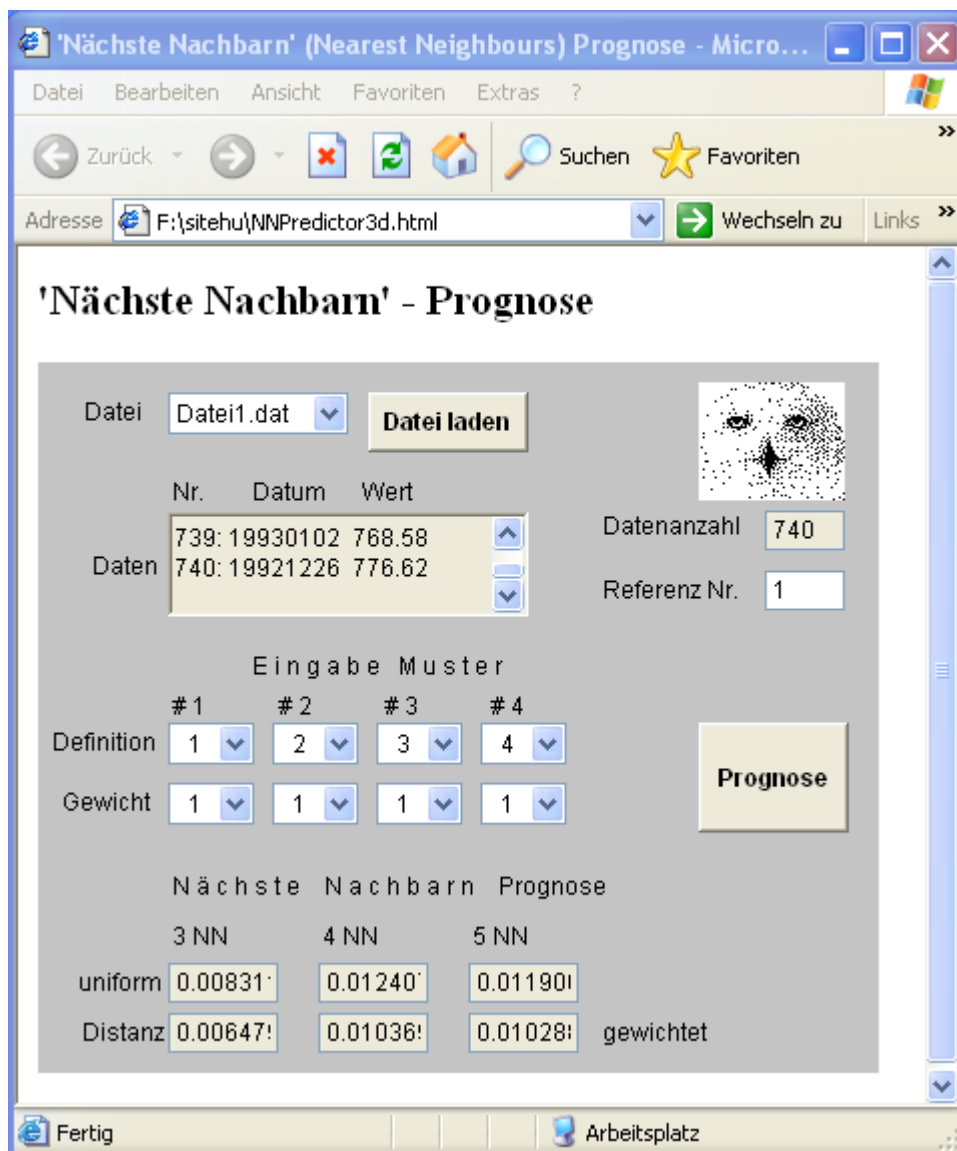


Abbildung 1: Aussehen des JAVA™-Applets nach dem Aufruf des Programms im Internet Explorer von Microsoft und dem Laden der Datei1.dat mit Daten des Österreichischen Aktienindex ATX. Anschließend wurde der Knopf <Prognose> angeklickt. Die Datei enthält insgesamt 740 Datensätze, d.h. Datumsangaben und Kurswerte. Man kann sie sich einzeln im Datenfenster ansehen, wenn man mit der Maus auf den rechten Balken geht, und blättert. Als Referenz wurde der Datensatz Nr. 2 gewählt, der dem Marktzustand am 16. Februar 2007 entspricht. Der Marktzustand wird hier definiert durch jeweils vier Marktbewegungen, # 1 bis # 4. Das hier gewählte Muster für einen Marktzustand besteht aus den Marktbewegungen in den vier voran gegangenen Wochen. Dies zeigt die Reihe rechts von der Beschriftung ‚Definition‘ an. Man hätte den Marktzustand auch durch ein anderes Muster definieren können. Wenn man auf den Pfeil klickt, der rechts neben den einzelnen Anzeigefeldern zu sehen ist, klappt eine Liste auf, aus der man unter verschiedenen Zahlen wählen kann. In den Feldern rechts von der Bezeichnung ‚Gewicht‘ ist die Gewichtung der einzelnen Eingabewerte angegeben. Die Zahlen entsprechen

jeweils den Gewichtungsfaktoren für den jeweiligen Eingabewert. Hier werden alle Angaben einheitlich gewichtet. Man kann zwischen den Faktoren 0,8 bis 1,2 wählen. Unterhalb der Überschrift ‚Nearest Neighbour Vorhersage‘ sieht man sechs Anzeigefelder, in denen die jeweilige Vorhersage ausgegeben wird. Diese Vorhersagen können sich durchaus widersprechen. In der oberen Reihe sind die Mittelwerte der Vorhersagen mit einheitlich gewichteten nächsten Nachbarn zu sehen, von links nach rechts sind es Mittelwerte aus den drei, vier oder fünf nächsten Nachbarn. Die in der unteren Reihe angezeigten Mittelwerte wurden erhalten, nachdem die Nächsten Nachbarn entsprechend ihrer Ähnlichkeit zum Referenzmuster gewichtet wurden.

## Besonderheiten von JAVA-Programmen

JAVA™ ist eine sehr moderne, rein objektorientierte, höhere Programmiersprache, die für das Internet entwickelt wurde. JAVA-Programme laufen daher üblicherweise auf Servern im Internet. Da das Internet für alle da ist, ob sie nun einen Windows-PC, einen Mac, eine Unix Workstation oder einen anderen Rechner verwenden, braucht man dafür Programme, die auf allen diesen Rechnern laufen. Genau für diese Aufgabe wurde JAVA entwickelt. Der JAVA-Programmtext wird von einem Compiler in einen Byte Code umgewandelt, der auf einem sogenannten 'virtuellen Computer' (englisch: virtual machine VM) läuft. Die so umgewandelten Programme laufen auf keinem einzigen realen Computersystem. Aber für jedes der oben erwähnten Computersysteme – und einige mehr - stellen die Entwickler von JAVA, also SUN Microsystems, Übersetzungsprogramme (Interpreter) zur Verfügung, mit denen der Byte Code in die jeweiligen, systemspezifischen bits und bytes übersetzt wird. Diese Interpreter sind in der JAVA-Laufzeitumgebung (JAVA runtime environment, JRE) enthalten, die man sich kostenlos aus dem Internet laden kann (~10MB), entpackt benötigt sie ~20 MB. JAVA kennt verschiedene Programmtypen. Es gibt einmal die eigenständigen Programme, für deren Ausführung man eine vollständige JRE benötigt, und Applets, kleine Programme, die im Web-Browser laufen können und aus Sicherheitsgründen bestimmten Einschränkungen unterliegen.

Ein JAVA™-Applet ist kein eigenständiges, ausführbares Programm. Um es laufen zu lassen, benötigt man einen der üblichen Web-Browser, wie z.B. den Internet Explorer von Microsoft™ (MSIE), aber dessen Funktion muss mit einem JAVA™-Plug-In erweitert sein. Wer noch einen älteren MSIE benutzt, kann das Programm wahrscheinlich sofort starten. Die älteren Versionen des MSIE, zumindest bis MSIE5.5 hatten einen JAVA JIT-Compiler bereits integriert, ebenso die älteren Netscape Navigatoren und ältere Versionen des OPERA Browsers. Inzwischen sind die Anbieter von Browsern dazu übergegangen, die JAVA-Fähigkeiten nur noch als Plug-In nachinstallieren zu lassen. Das macht zwar etwas Mühe, doch es bewirkt, dass die Applets zumindest auf der gleichen Plattform (Betriebssystem) gleich aussehen. Anstelle eines Web-Browsers kann man für Applets auch den Appletviewer von SUN™ Microsystems verwenden, falls sich das Applet auf dem Rechner des Benutzers befindet. Dieser wird jedoch nur mit dem Entwicklerwerkzeug für JAVA-Programmierer ausgeliefert, das zwar ebenfalls kostenlos aus dem Internet geladen werden kann, aber etwa 300 MB Speicherplatz erfordert.

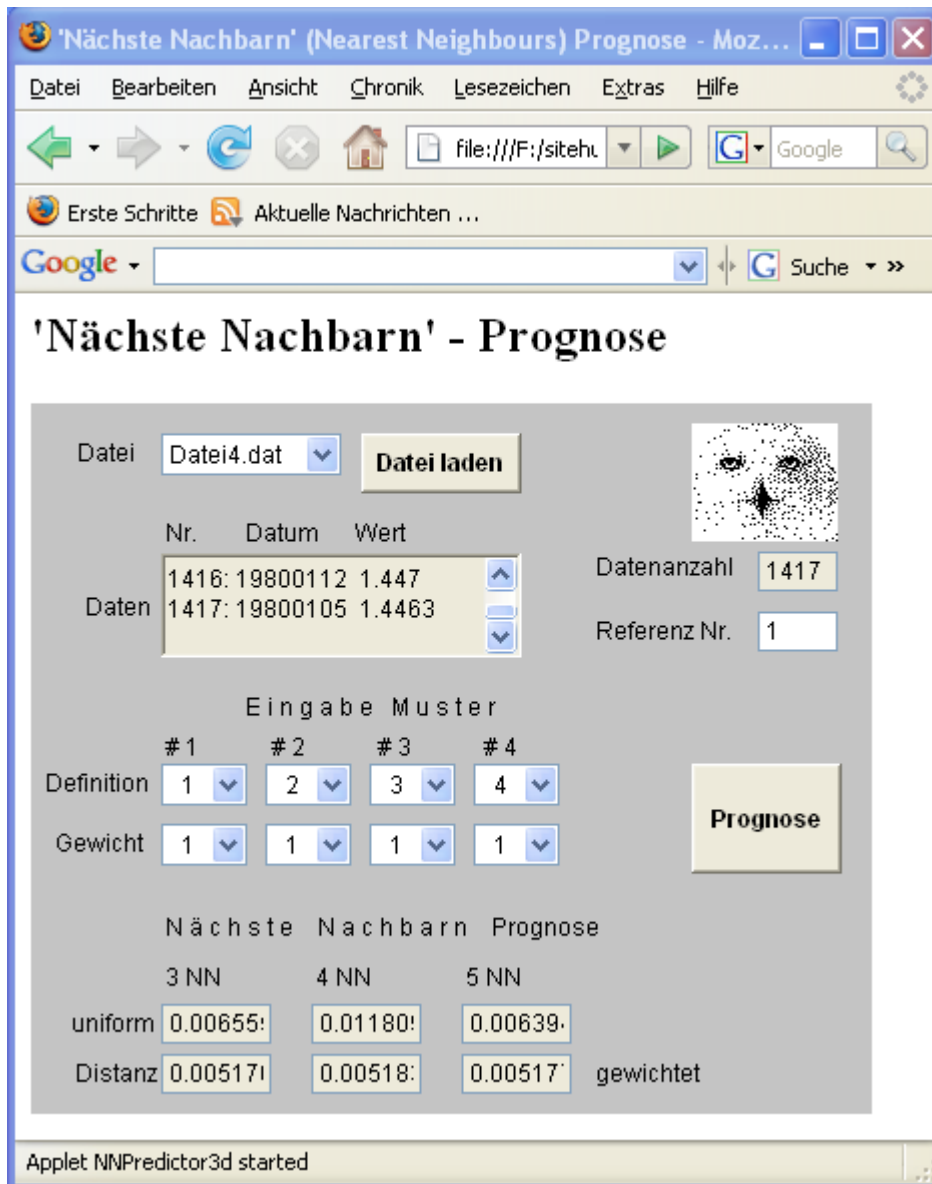


Abbildung 2: Aussehen des JAVA™-Applets nach dem Aufruf des Programms im Mozilla Firefox und dem Laden der Datei4.dat mit Daten der Wechselkurse des Euro in US-\$. Anschließend wurde ebenfalls der Knopf <Prognose> gedrückt.

Die Programme können etwas unterschiedlich aussehen, je nachdem, unter welchem Betriebssystem man arbeitet und welchen Browser man verwendet, oder ob man es im Appletviewer laufen lässt. Bildschirmausdrucke der verschiedenen Programmansichten finden Sie auf den folgenden Seiten.

### Daten

Die Daten müssen sich in Dateien mit den vorgegebenen Namen befinden. Ich füge fünf Dateien bei, in denen Wochenschlusskurse von Aktienindizes und Devisen gespeichert sind. Im einzelnen sind dies Daten des österreichischen ATX, des deutschen DAX und des schweizerischen SMI, dazu die Wechselkurse des US-\$ zum Euro und zum Schweizer Franken. Die Datendateien enthalten neben den Kursangaben auch das jeweilige Datum. Oben in der Datei ist auch die Anzahl der Daten angegeben. Es sind reine ASCII-Dateien, man kann sie sich in jedem Texteditor ansehen. Die mitgegebenen Dateien kann man, falls man dies möchte, durch andere ersetzen, sofern diese das gleiche Format besitzen, aber die Namen sind fest kodiert und dürfen nicht geändert werden. Die Dateien müssen sich auch im gleichen Verzeichnis befinden, wie das Applet, andernfalls werden sie nicht gefunden.

### Daten laden

Wenn man das Programm aufgerufen hat, geht man auf die Dateienliste und kann sich eine der Dateien durch Anklicken auswählen. Dann drückt man den Knopf <Daten laden> und man sieht im Datenfenster, wie die Daten geladen werden. In einem Sichtfenster wird die Datenanzahl ausgegeben.

### Referenznummer

Diese gibt die Nummer des eingelesenen Datensatzes an, von dem aus man eine Woche in die Zukunft blicken möchte. Voreingestellt ist die 1. Das bedeutet, dass man ausgehend vom aktuellsten Datensatz die Vorhersage machen möchte. Das Feld mit der Referenznummer ist editierbar. Man kann also eine beliebige Zahl eingeben. Indem man nacheinander aufsteigende Zahlen eingibt, kann man überprüfen, wie sich die Methode in der Vergangenheit bewährt hat. Man kann sich auch markante Punkte in der Vergangenheit suchen, z.B. irgendwelche Kursgipfel und prüfen, ob die Methode mit den jeweils gewählten Parametern bei der Vorhersage richtig gelegen hätte.

### Datenmuster

Ein Marktzustand wird durch ein Datenmuster aus jeweils vier Daten charakterisiert. Wie dieses Muster definiert wird, kann man über die Auswahllisten in der oberen Reihe festlegen. Im ersten Feld des Musters (#1) kann man ganzzahlige Werte von 1 bis 4 anklicken, in Feld #2 sind es Werte von 2 bis 7, in Feld #3 Werte von 3 bis 8 und in Feld #4 kann man Werte von 4 bis 9 auswählen.

In der unteren Reihe kann man für jeden Eingabewert noch eine individuelle Gewichtung auswählen. Wenn man die Pfeile rechts neben den Anzeigefeldern anklickt, klappen Listen auf, in denen man zwischen Werte zwischen 0,8 und 1,2 wählen kann. Wenn man also meint, der jeweils aktuellste Wert sei wichtiger, als der letzte, dann kann man ihn stärker gewichten. Wenn man eine robustere Bewertung möchte, sollte man die beiden mittleren Werte stärker gewichten, als den jüngsten und den ältesten Eingabewert.

### Prognose

Der Prognose-Knopf ist unwirksam, solange keine Daten geladen wurden. Wenn eine Datei geladen ist, und man klickt den Knopf an, dann werden die voreingestellten Parameter verwendet. Sobald man einen Parameter geändert hat, kann man den Knopf anklicken und man erhält wieder eine neue Vorhersage auf der Basis der aktuellen Parameter.

Die Vorhersagewerte werden in den unteren sechs Feldern angezeigt, in der oberen Reihe auf der Basis einheitlich gewichteter nächster Nachbarn, in der unteren auf der Basis abstandgewichteter nächster Nachbarn. Die angezeigten Beträge sind in der Regel kleiner als 0,1 und tragen entweder kein Vorzeichen oder ihnen ist ein Minuszeichen vorangestellt. Die absoluten Zahlenwerte sollten Sie nicht kümmern, denn sie sind allenfalls Anhaltspunkte. Eigentlich geht es nur um die zu erwartende Markttrichtung. Dafür ist allein wichtig, ob positive oder negative Bewegungen angezeigt werden.

Gelegentlich werden auch Werte angezeigt, die Ziffern von 1 bis 9 vor dem Komma tragen. Diese Anzeigen sind keine Programmfehler. Tatsächlich sind es dann Werte mit sehr kleinen Beträgen, die der Computer in Exponentialschreibweise angibt, weil dies eigentlich übersichtlicher ist. Wenn beispielsweise sieben oder mehr Nullen hinter dem Komma folgen, kann man schnell mit den Dimensionen durcheinander geraten. Der Computer schreibt dann, sagen wir 9,87654321E-7, das heißt, die erste Zahl ungleich Null steht erst an der 7. Stelle hinter dem Komma. Die Anzeigefelder sind jedoch aus Gründen der Übersichtlichkeit schmal gehalten, so dass für den Betrachter der hintere Teil der Exponentenangabe nicht zu sehen ist.

## Arbeiten mit dem Programm

Bevor Sie mit dem JAVA-Applet arbeiten können, müssen Sie Ihren Web-Browser JAVA-fähig machen. JAVA ist eine sehr sichere Programmiersprache, die auf Ihrem Computer nichts verändern kann. Sie müssen also keine Angst davor haben. Laden Sie sich das JAVA-Plugin für Ihren Browser aus dem Internet. Wenn Sie versuchen, ein JAVA-Applet aufzurufen und ihrem Browser fehlt das notwendige Plugin, dann fragt eine Message-Box, ob Sie das Plugin nachinstallieren wollen. Wenn Sie online sind, müssen Sie dazu nur den angezeigten download-Knopf anklicken. Das Installieren des Plugins geschieht in der Regel automatisch. Falls das nicht klappt, z.B. weil das vorgegebene Laufwerk nicht genügend freien Speicherplatz besitzt, wird das Installationsprogramm auf ihrem Desktop abgelegt. Sie müssen es dann aufrufen und im Dialog geführt installieren.

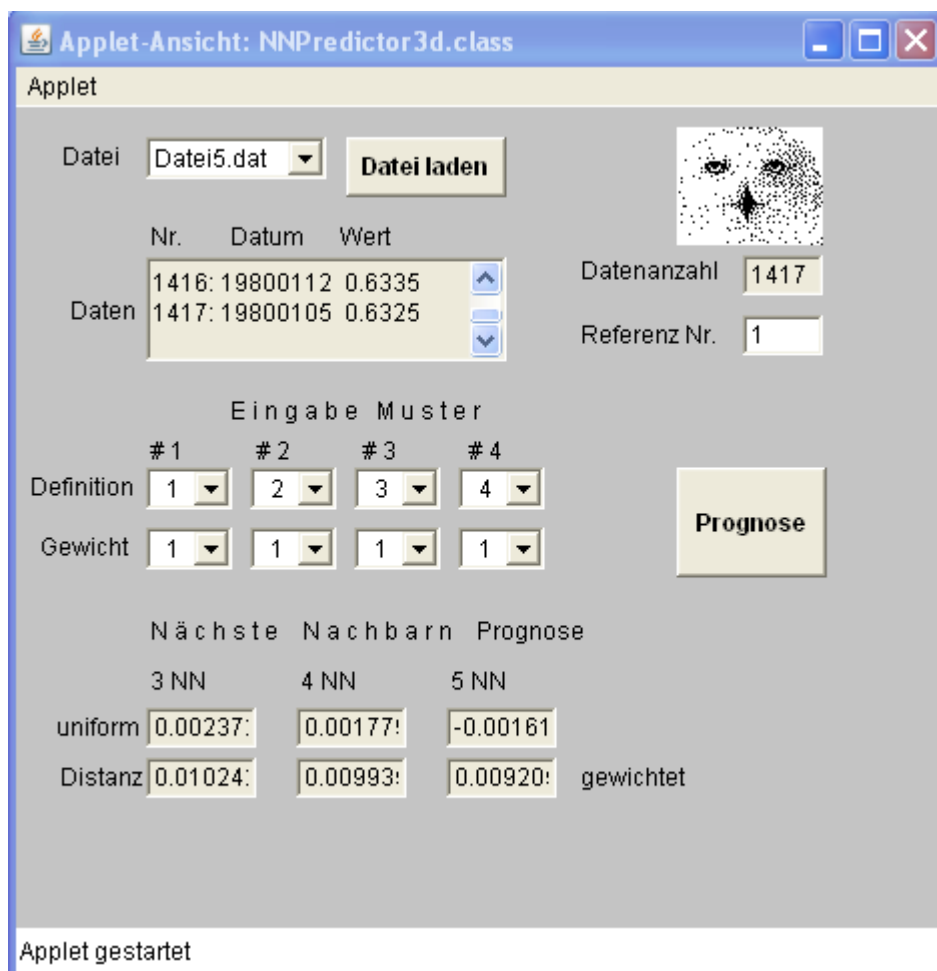


Abbildung 3: Das Programm im Appletviewer von SUN™ Microsystems – Der Appletviewer kann nur Applets ausführen. Das Erscheinungsbild ist ansonsten wie bei Abbildung 1 besprochen. Es wurde hier jedoch die Datei5.dat aufgerufen, sie enthält Wechselkurse des Schweizer Franken in US- $\text{\$}$ .

Mit diesem Applet kann man, sofern genügend Zeit dafür übrig ist, beliebig viele Parameter ausprobieren, bis man geeignete gefunden hat, die auch über einen längeren Zeitraum überwiegend korrekte Vorhersagen liefert. Auf jeden Fall geht diese sehr viel schneller, als mit einem Kalkulationsblatt. Die Prognosen sollten in Zeiten steigender Kurse und in Zeiten fallender Kurse etwa gleich zuverlässig sein. Man muss sich nicht für eine bestimmte Anzahl nächster Nachbarn entscheiden, oder ob man einheitlich gewichtete oder abstandsgewichtete nächste Nachbarn verwendet. Da immer alle sechs Varianten angezeigt werden, wäre es auch möglich, nach Mehrheitsvotum zu entscheiden. Dafür muss man nicht alle sechs Anzeigen heranziehen. Man kann sich auch auf die Mehrzahl von drei, vier oder fünf Vorhersagen stützen.