

ЭКСПЕРТНАЯ СИСТЕМА ДЛЯ ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЯ ВСХОЖЕСТИ ПШЕНИЦЫ

О.В. Лукьянычева, С.П. Пронин

В статье описана экспертная система для определения показателя всхожести зерен пшеницы, основанная на анализе биоэлектрического сигнала. Экспертная система использует байесовский вероятностный подход и позволяет ускорить процесс определения показателя всхожести в 10-12 раз.

Ключевые слова: биоэлектрический сигнал, экспертная система, теорема Байеса.

Семена, предназначенные для реализации в пределах России, а также поставки в региональные и федеральные фонды, подлежат сертификации по показателям, удостоверяющим их сортовые и посевные качества. К посевным качествам относятся: чистота семян, всхожесть, влажность, крупность, зараженность болезнями и заселённость вредителями и ряд других. Наиболее длительным анализом является определение показателя всхожести. В соответствии с ГОСТ 12038-84 для семян пшеницы срок определения составляет 7-8 суток. В работах [1,2] показано, что для определения показателя всхожести пшеницы возможно использование анализа биоэлектрического сигнала зерен, что позволяет значительно ускорить процесс определения всхожести. Для реализации данного подхода разработан аппаратно-программный комплекс, включающий экспериментальную установку и программные средства для предварительной обработки результатов и подготовки данных для информационной системы по определению показателя всхожести зерен.

Задача определения всхожести зерен имеет статистический характер и обладает ошибочностью, неполнотой, неоднозначностью, противоречивостью, как в исходных данных, так и в плане знаний о решаемой задаче. Неточность экспериментальных результатов связана как с объективными причинами: различием энергии прорастания у каждого отдельного зерна, аппаратными ограничениями экспериментальной установки, - так и с субъективными причинами: особенностями объекта исследования и, как результат, невозможностью получения идентичности проведения опытов, временными ограничениями проведения замеров и так далее. Не надежность знаний связана с отсутствием формальных процедур получения точных

данных, вероятностной природой поступающих данных, недостаточной математической и логической обоснованностью правил определения всхожести.

В связи с этими особенностями, в информационных системах анализа и принятия решения используются вероятностные оценки тех или иных знаний, как в части данных, так и правил вывода. В исследованиях по искусственному интеллекту подобными неформализованными задачами занимается направление, получившее название "экспертные системы".

Принятие решения носит экспертный (субъективный) характер, свойственный задачам, решаемым в нечетких предметных областях. Статические доопределяющие ЭС применяются для решения задач с такими неполными, неточными данными и знаниями. Вероятностный, статистический принцип, заложенный в процесс обработки экспериментальных данных, свидетельствует, что наиболее подходящим методом принятия решения в исходных условиях является вероятностный подход на основе теоремы Байеса.

Процесс определения показателя всхожести зерен пшеницы занимает 14 часов.

Зёрна помещаются в экспериментальную установку и закладываются в ячейки поролоновых форм, которые ставят в пластиковые лотки по две формы в лоток, и заливают дистиллированной водой. Лотки помещают в полиэтиленовые пакеты, для того чтобы предотвратить выпаривание воды из поролона. Зёрна набухают в течение 12 часов при заданной температуре 21°C в экспериментальной установке [3]. После этого производится измерение параметров биоэлектрических сигналов, у которых фиксируется девять наиболее информативных показателей [4]. Для формирования базы знаний результаты анализа эталонных зерен, всхо-

ЭКСПЕРТНАЯ СИСТЕМА ДЛЯ ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЯ ВСХОЖЕСТИ ПШЕНИЦЫ

жесть которых определена в соответствии с ГОСТ, подвергаются статистической обработке. В базу знаний помещаются числовые значения показателей биоэлектрического сигнала и интервальные характеристики этих показателей. Для преодоления неоднозначности в экспериментальных данных база знаний дополнена экспертными оценками значимости показателей биоэлектрического сигнала.

В ЭС к решателю, использующему выбранный метод, подключается база знаний и рабочее поле. В рабочее поле из аппаратно-программного комплекса предварительной подготовки загружаются узловые параметры исследуемых графиков биоэлектрических по-

тенциалов. Таким образом, в рабочем поле находятся значения результатов из 4 поролоновых форм - 4 пробы. Каждая проба содержит 9 показателей для одного зерна, причем в пробе не менее 30 зерен.

Принятие решения возможно двумя способами: экспресс-анализ и углубленный анализ.

Диаграмма экспресс-анализа представлена на рисунке 1.

В режиме экспресс-анализа предварительно вычисляются выборочные средние для каждого из 9 показателей исследуемых проб и формируются усредненные наборы значений параметров по каждой пробе.

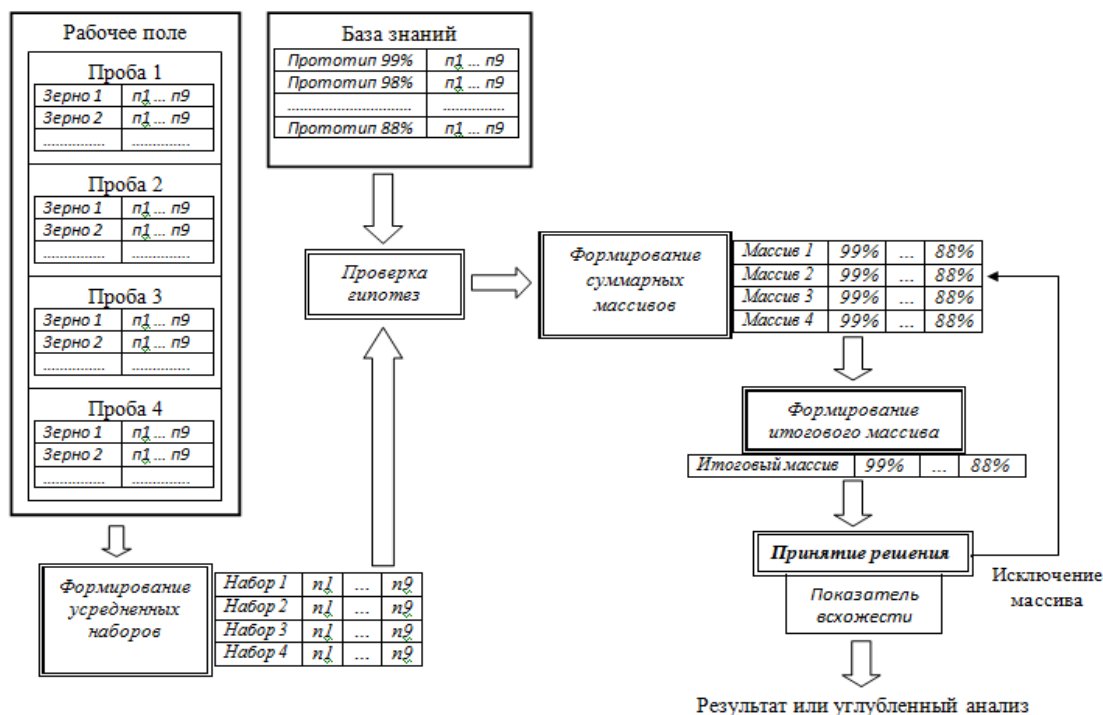


Рисунок 1 - Диаграмма экспресс-анализа

Для каждого усредненного значения параметра пробы проверяется гипотеза принадлежности к определенной всхожести (нулевая гипотеза). Если проверяемая величина попадает в доверительный интервал, нулевая гипотеза не отклоняется, поскольку исследуемый параметр не является необычным, и оценке достоверности нулевой гипотезы присваивается значение единицы. С другой стороны, если проверяемая величина не попадает в доверительный интервал, нулевая гипотеза отклоняется, поскольку исследуемый параметр является экстремальным, и оценке достоверности нулевой гипотезы при-

сваивается значение нуля. По результатам проверки гипотез формируется оценка доверительной вероятности с учетом коэффициентов степени влияния (веса) отдельно по каждому классу всхожести. Таким образом, для каждой пробы формируется суммарный массив оценки всхожести.

По четырем пробам вычисляется в соответствии с теоремой Байеса вероятность по каждой всхожести:

$$p(H_i | E_1 E_2 \dots E_n) = \frac{p(E_1 E_2 \dots E_n | H_i) \times p(H_i)}{\sum_{k=1}^m p(E_1 E_2 \dots E_n | H_k) \times p(H_k)} ,$$

где H_i – i -ая гипотеза всхожести, m – количество гипотез, равное 8, E_j – j -ое свидетельство (проба) всхожести, n – количество проб, равное 4.

Возможность предположения об условной независимости экспериментов отдельных проб друг от друга позволяет упростить правило Байеса:

$$p(H_i|E_1 E_2 \dots E_n) = \frac{p(E_1|H_i) \times p(E_2|H_i) \times \dots \times p(E_n|H_i) \times p(H_i)}{\sum_{k=1}^m p(E_1|H_k) \times p(E_2|H_k) \times \dots \times p(E_n|H_k) \times p(H_k)}$$

Вероятности принадлежности испытуемых проб к любой всхожести одинаковы, то есть $p(H_i) = 1/8$.

По максимальному значению в итоговом массиве определяется доверительная вероятность принадлежности испытуемого зерна к определенной всхожести (таблица 1).

Таблица 1 – Пример формирования итогового массива оценки принадлежности 4 проб к определенным классам всхожести

Всхожесть	88	89	90	92	95	97	98	99
Проба 1	0,2242	0,3402	0,7617	0,4586	0,3452	0,4664	0,2383	0,2223
Проба 2	0,4647	0,5768	0,6485	0,2940	0,4647	0,4141	0,1160	0,1769
Проба 3	0,5245	0,2942	0,6416	0,2890	0,0648	0,0648	0,1769	0,0648
Проба 4	0,4597	0,2303	0,5768	0,2242	0	0	0,1121	0,1160
Итоговый массив	0,1088	0,0576	0,7920	0,0378	0	0	0,0024	0,0013

В качестве результата принимают всхожесть, у которой наблюдается максимальное значение итогового показателя доверительной вероятности.

Если результат неудовлетворительный, то при проведении анализа по четырем пробам пользователь (эксперт) может исключить из анализа любую из 4 проб и вычислить оценку показателя всхожести повторно.

Если результат, с точки зрения пользователя, неудовлетворительный и в этом случае, то пользователь может перейти к углубленному анализу.

В режиме экспресс-анализа пользователь имеет возможность просмотреть суммарные массивы оценок всхожести, оценки доверительной вероятности у каждой партии и проанализировать причину неудовлетворительного результата.

Основной особенностью углубленного анализа является анализ показателей по каждому образцу зерна проверяемой партии. Это позволяет выяснить причины неудовлетворительного результата при экспресс-анализе и в пределах допустимых ограничений провести необходимые действия для получения более доверительного результата. Диаграмма углубленного анализа представлена на рисунке 2.

Набор из девяти показателей каждого графика в каждой пробе сравнивается с прототипами из базы знаний - проверяются статистические гипотезы принадлежности ко всем классам всхожести (прототипам) из базы знаний с учетом весовых коэффициентов

(степени влияния) и определяются оценки доверительной вероятности для каждого зерна по каждому классу всхожести (промежуточный массив).

При этом математическое ожидание μ_0 и стандартное отклонение σ равны соответствующим значениям из прототипов базы знаний. Уровень значимости α равен 0,05.

Статистика g вычисляется по формуле:

$$g = \frac{|\bar{x} - \mu_0|}{\frac{\sigma}{\sqrt{N}}}$$

где \bar{x} - выборочное среднее, N - количество проверяемых зерен.

Распределение статистики g при справедливой гипотезе является нормальным распределением. Критическая область отклонения гипотезы определяется в соответствии с соотношением $g > z_{1-\alpha}$, где $z_{1-\alpha}$ - квантиль порядка $(1 - \alpha)$, то есть $z_{1-\alpha} = 1,644854$.

Если неравенство выполняется, то нет оснований отвергать гипотезу, и оценке достоверности гипотезы присваивается значение единицы. Если неравенство не выполняется (критическая область), то есть основания для того, чтобы отвергнуть гипотезу (это маловероятный результат), и оценке достоверности гипотезы присваивается значение нуля. В режиме углубленного анализа $N=1$, \bar{x} - зна-

ЭКСПЕРТНАЯ СИСТЕМА ДЛЯ ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЯ ВСХОЖЕСТИ ПШЕНИЦЫ

чение проверяемого параметра. В базе знаний предусмотрена возможность вычисления стандартного отклонения по доверительному интервалу. Для этого добавлен параметр –

значение величины выборки (генеральной совокупности) для получения прототипов.

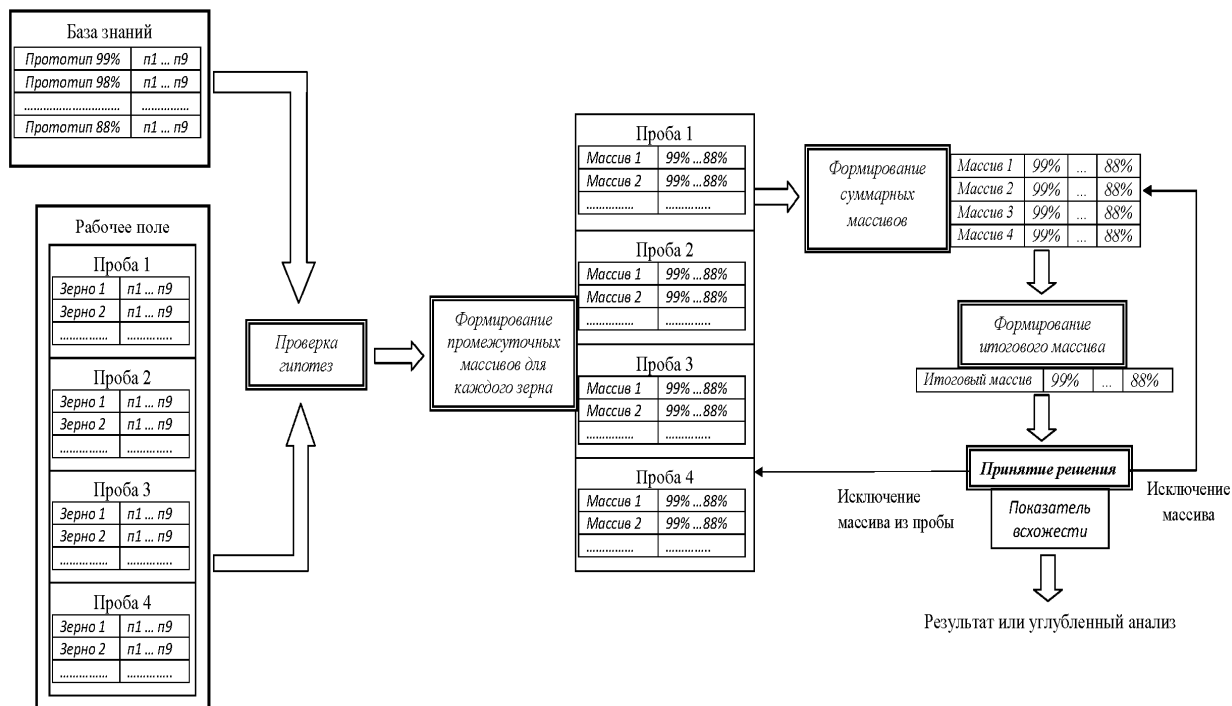


Рисунок 2 - Диаграмма углубленного анализа

По результатам оценивания гипотез для каждого образца (зерна) формируется промежуточный массив доверительных вероятностей принадлежности образца ко всем классам всхожести.

Суммарный массив пробы содержит значения математических ожиданий, полученных по всем промежуточным массивам отдельных образцов.

Вычисляется математическое ожидание:

$$M(x) = \sum_{i=1}^n p_i x_i,$$

где n – количество зерен в пробе, p_i – вероятность события x_i для каждой всхожести.

Вероятность принадлежности каждого зерна к определенной всхожести равновероятна и равна $p_i = 1/n$.

Затем, как и в экспресс-анализе, сформированные суммарные массивы для каждой из четырех проб подвергаются байесовскому оцениванию, и в результате формируется итоговый массив, представляющий собой обобщенные значения оценок принадлежности всех проб по каждому классу всхожести.

По максимальному значению показателя в итоговом массиве выбирается гипотеза с максимальным доверием, которая принимается как результат оценки принадлежности испытуемого зерна к определенному классу всхожести.

Если отличие предлагаемого решения незначительно по отношению к другим альтернативным гипотезам, то возможна оценка по трем пробам.

Если результат не удовлетворяет пользователя (эксперта), то он может посмотреть все промежуточные расчеты по отдельным партиям и удалить самые ненадежные результаты в количестве, определяемом ограничением стандартного нормального распределения. После этого следует провести анализ повторно.

Если и в этом случае результат не кажется пользователю (эксперту) надежным, то необходимо повторить эксперимент.

При настройке и тестировании экспертной системы эксперт имеет возможность настраивать правила формирования решения, анализируя результаты решения. Кроме того, эксперт имеет возможность пополнять базу

знаний за счет новых экспериментальных данных.

Специализированная экспертная система «Определение показателя всхожести зерна» показала эффективность применения предложенного подхода для определения всхожести зерна пшеницы мягких сортов. Разработанная экспертная система является универсальной и может применяться для определения всхожести других зерен за счет переключения базы знаний и, если необходимо, таблицы оценок экспертов.

СПИСОК ЛИТЕРАТУРЫ

1. Матлаев, А.Г. Метод и средство контроля всхожести семян пшеницы /А. Г. Матлаев, С. П. Пронин // Естественные и технические науки. – 2009. – № 3. – С. 308-311.

2. Мерченко, Н.Н. Разработка метода контроля всхожести зерен пшеницы по мембранному потенциалу/ Н.Н Мерченко., С.П. Пронин, А.Г. Зрюмова // Вестник Алтайского государственного аграрного университета. – 2013. – № 10 (108). – С. 103-106.

3. Лукоянычева, О.В. Исследование электрических сигналов в зёрнах пшеницы с различной всхожестью и разработка рекомендаций по построению экспертной системы / О.В. Лукоянычева, С.П. Пронин // Вестник Алтайского Государственного Аграрного Университета. – 2014. - №1(111). – С. 109-114.

4. Лукоянычева, О.В. Автоматизация процесса определения показателя всхожести зерен пшеницы с использованием биоэлектрических сигналов / О.В. Лукоянычева, С.П. Пронин // Известия АГУ. – 2014. – № 1-2(81). – С. 108-113.

Лукоянычева О.В. – магистрант кафедры ИТ, АлтГТУ Им. И. И. Ползунова,
тел. 8(3852)-29-07-96,
lukoya@bk.ru.

Пронин С.П. – д.т.н., профессор, заведующий кафедрой ИТ, АлтГТУ Им. И. И. Ползунова,
тел. 8(3852)-29-07-96,
sppronin@mail.ru.